



**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicant: Caius Rommens *et al.*

Title: PRECISE BREEDING

Appl. No.: 10/607,538

Filing Date: 6/27/2003

Examiner: David T. Fox

Art Unit: 1638

**DECLARATION UNDER 37 C.F.R. § 1.132**

Commissioner for Patents  
PO Box 1450  
Alexandria, Virginia 22313-1450

Sir:

1. I, Caius Rommens, have been employed as Director R&D at J. R. SIMPLOT COMPANY (Boise, Idaho), since 2000.

2. I have doctorate in Plant Biology from the Free University of Amsterdam. I have worked extensively in agricultural genetics throughout my professional career. I attach a Curriculum Vitae evidencing my research in this field. See Exhibit 1. I authored the peer-reviewed article Rommens *et al.*, "Plant-derived transfer DNAs," *Plant Physiol.* 139: 1338-49, 2005 (Exhibit 2), for example, which describes our presently claimed invention.

3. In this regard, I am the project director and co-inventor of the present application, USSN 10/607,538, which discloses and claims key SIMPLOT technologies. Our "all-native" green biotechnology approach to genetic engineering of plants is novel and is generating much excitement and interest in the agricultural community. Our goal, which is evidenced throughout our application, is to transform plants without necessarily incorporating bacterial or viral or non-plant DNA into the recipient plant's genome.

4. We have successfully achieved this goal as evidenced by the specification and examples disclosed in our application. Thus, we prove, by way of our originally-filed specification, that it is possible to transform a plant with a desirable gene or polynucleotide without necessarily incorporating any "foreign" DNA into the plant genome. This is a **major breakthrough** on many different levels, especially those dominating the agricultural and genetic engineering communities. The purpose of my declaration is to further explain and clarify our claimed invention, particularly with respect to present claim 3 under examination.

5. In this respect, I have read the Office Action dated May 23, 2006, and understand the Office's position regarding enablement as follows:

[G]iven the claim breadth, unpredictability, and lack of guidance . . . undue experimentation would have been required . . . to identify and isolate a multitude of non-exemplified P-DNAs or "transfer-DNAs" of a multitude of lengths and sequences, from a multitude of plant species and genomic regions [as well as to functionally] evaluate said non-exemplified sequences (page 13, lines 4-13).

I offer the following commentary and evidence to respectfully rebut this conclusion. Specifically, I clarify how the application provides explicit guidance for identifying and isolating border-like sequences, and how these border-like sequences can be readily tested to determine their respective ability to support plant transformation. I will explain and prove that **the methods disclosed in the application most certainly do yield predictable results.**

6. Two particularly interesting methodological aspects of our invention, which are extensively disclosed in the present *Precise Breeding* application, include (1) **the identification and isolation of plant-derived sequences** that resemble, but are not identical to, the nucleotide sequences of *Agrobacterium* T-DNA borders, and (2) **the use of such border-like sequence for plant transformation.**

7. It is clear from the studies disclosed in our application and our subsequent use of those disclosed methods, that **predictability is not a questionable issue.** As I elaborate in this declaration, our disclosed methods for identifying and isolating sequences that conform to nucleotide composition requirements of transformation borders is predictable. Secondly, our functional assay is easy to perform and does not involve any undue experimentation. Simply put, our disclosed functional assay predictably identifies which border-like sequence is effective in transforming a plant.

8. A valuable outcome of our extensive and accurate analyses of the nucleotide sequences that we isolated, identified, and tested for functional activity led us to define a P-DNA border-like sequence by all of the following criteria:

(i) it consists of 25 nucleotides and displays between 52 and 96% identity to a functional *Agrobacterium* T-DNA border;

(ii) it can (a) be identified by searching databases for genomic and/or cDNA sequences of plants with any query sequence that complies to the *Agrobacterium* T-DNA border motif YGRYAGGATATATWSNVBKGTAAWY, or (b) be isolated by employing polymerase chain reactions using primers that anneal to T-DNA border sequences;

(iii) it enables the transformation of a plant cell with DNA from a plasmid that contains the P-DNA border-like sequence.

I discuss this enabling disclosure in more detail below.

9. Our specification discloses particular methods for identifying and isolating border-like sequences. For example, we relate that one may **(1) search DNA databases** and **(2) employ the polymerase chain reaction (PCR)** to identify sequences and actual DNA molecules that satisfy our inventive “border-like” criteria. See the following paragraphs.

10. With respect to searching DNA databases for border-like sequences, the specification provides:

(A) **Examples of databases** that can be searched for sequences that share identity with a non-T-DNA border sequence and/or which meet the requirements of our border-like consensus sequence (see paragraph 326), e.g., (i) The National Center For Biotechnology Information, (ii) The Institute for Genomic Research, and (iii) SANGER; and

(B) **Search tools** for performing searches, e.g., (i) Motif Alignment and Search Tool (Bailey and Gribskov, *Bioinformatics* 14: 48-54, 1998), and (ii) advanced BLASTN (“penalty for nucleotide mismatch” = -1; “expect” =  $10^5$ ; Altschul et al., *Nucleic Acids Res* 25: 3389-3402, 1997) (see paragraph 327). See, for example, Exhibit 3.

11. With respect to using PCR to physically amplify and isolate border-like sequences, the specification provides “one-step” and “two-step” amplification protocols.

(A) **One-Step PCR border-like DNA isolation method**

(a) The present specification describes how polynucleotides that comprise border-like sequences can be amplified from plant DNA using PCR primers that are designed to anneal to border or border-like sequences. See paragraph 320 of the specification.

(b) The common term “anneal” is readily understood by any molecular biologist. An annealing primer may represent either the exact inverse complement of a target sequence or it may be somewhat dissimilar in sequence to the target. Nonetheless, whether the primer is or is not an exact complement of the target, it should be able to hybridize with its target sequence at a set “annealing” temperature. Paragraph 47 explains that primers may be degenerate.

(c) Specific examples of primers are shown in paragraph 165 (CAGGATATATNNNNKGTAAAC), paragraph 321 (TGR CAGGATATATNVNDNTGTAAAC) and paragraph 330 (YGR CAGGATATAT). A variety of such primers that were designed to anneal to border sequences were used to isolate border and P-DNA sequences from potato, e.g., as depicted in SEQ ID NO. 1. See paragraph 320.

(B) **Two-Step PCR border-like DNA isolation method**

Plant DNA fragments with an unknown sequence were ligated to a known sequence. By using the resulting DNA as a template, it was possible to use (i) a single primer designed to anneal to a border sequence (the border primer) and (ii) a primer that anneals to the known sequence (the anchor primer) to perform a two-step PCR. See paragraph 321-330. An example of the method also is shown in Exhibit 4.

**Step 1** entails (a) digesting total DNA with a restriction enzyme such as *Sau*III; (b) ligating the restriction fragments to a known polynucleotide, such as to the 192 bp *Bam*HI-*Eco*RV fragment of pBR322; (c) performing a first PCR with a border primer (e.g., YGRCAGGATATATNNNNKGTAAAC) and an anchor primer (e.g., GACCACACCCGTCCTGTG); and (d) ligating a PCR product to a second known polynucleotide such as pGEM-T;

**Step 2** entails (a) using the ligated DNA for a second PCR with a border primer and second anchor primer such as SP6 or T7; (b) cloning and sequencing the products; (c) using the sequenced DNAs to design four primers, and performing an inverse PCR on plant DNA that was digested with *Sau*III and self-ligated; and (d) cloning and sequencing the resultant amplified products to obtain actual border-like sequences.

12. These disclosed database-searching and PCR-based techniques ultimately produce a collection of plant-specific DNA sequences that can serve as potential T-DNA border-like elements. In this regard, we make clear in our application that a functional plant-specific border-like sequence does not have a nucleotide sequence that is identical to a conventional T-DNA border sequence, e.g., a left border or a right border. Accordingly, our reproducible methods help to uncover genetic elements that are useful in transferring one polynucleotide to another, e.g., from an *Agrobacterium* vector to a plant genome.

13. We teach in our application how to test the functionality of any one of those DNA sequences. See paragraph 345, as well as paragraph 20 below. Collectively, therefore, our methods provide DNA sequences that are (a) from a plant and (b) functional in transferring polynucleotides into, for instance, a plant genome. These DNA sequences therefore can be used to replace conventional T-DNA left and right border sequences in any given transfer cassette.

14. The functional DNA sequences can be compared with each other to determine whether there exists any commonality in sequence or nucleotide residues that can be used to characterize the sequences as a whole. This is precisely what we did in the present application. We analyzed the nucleotide sequences of the isolated genetic elements and derived a border motif nucleotide sequence. The border motif simply provides guidance concerning those nucleotide residues and features that we typically identified from our functional border-like sequence analysis. The consensus sequence does not signify any particular sequence or sequences that a potential DNA fragment must comport with in order to be characterized as a plant border-like sequence.

15. From our sequence searches of the disclosed databases and from our technical expertise in amplifying borders from plant genomic DNA, therefore, we identified and collated numerous border-like sequences. Our analysis of various border sequence comparisons (see paragraph 262) helped us to create the T-DNA **border motif** disclosed in Table 2 of the application, namely YGRYAGGATATATWSNVBKGTAAWY, where Y = C or T, R = A or G, W = A or T, S = C or G, K = G or T, V = A, C, or G, B = C, G, or T, and N = any nucleotide.

16. One particular sequence of this border motif is CGACAGGATATATTTGTTGGTAATC. This sequence was used to conduct an electronic BLAST search of the NCBI GenBank database with this sequence and identified sequences

that shared sequence identity. From those comparisons, we identified plants, e.g., **pepper**, **tomato**, and **maize**, which actually contained corresponding endogenous sequences.

17. Within that border motif sequence, we identified a **central core region** ANGATNTATN<sub>6</sub>GT that can be important in predicting functional activity in mediating plant transformation. See SEQ ID NO. 93 and paragraph 219.

18. Hence, we (1) created, from our analysis of numerous border sequences, a border motif sequence; (2) deduced polynucleotide sequences conforming to that consensus; (3) identified endogenous and native sequences that shared significant sequence identity with our deduced sequence, (4) designed PCR primers based on consensus sequences that can be used to isolate functional border-like sequences for any plant species, and (5) identified a central core region that often represents a good predictor for functional activity of the native sequences.

19. In fact, by way of these disclosed methods, **we successfully identified forty-one new border-like sequences**. See Rommens *et al. supra*.

20. In addition to those border-like sequence isolation and identification methods, we also disclosed in our specification a method to test the functionality of newly-identified border-like sequences. We performed this exact method to demonstrate that the potato P-DNA of SEQ ID NO. 1 could successfully transform tobacco. See paragraph 345. To summarize, our **specification teaches the following functional assay**:

(a) insert the border-like sequence and a neomycin phosphotransferase (*nptII*) gene into the linearized DNA of a plasmid that lacks T-DNA border sequences, such as pSIM100-OD-IPT; and then introduce that plasmid into an Agrobacterium strain. See paragraphs 31, 341, and 344.

(b) infect an explant of, e.g., tobacco, with the Agrobacterium. See paragraph 345.

(c) identify the subsequent formation of kanamycin resistant calli. Explants that contain the *nptII* gene will be kanamycin resistant and therefore identifiable as successful transformation events. The number of individual calli that are identified serves as a general indicator of the transformation efficiency of the border-like element. See paragraph 345.

21. We also used the traditional tobacco model to test the transformation capabilities of numerous border-like sequences that we had identified and isolated from **Arabidopsis, potato, tomato, barrel medic, alfalfa, barley, canola, corn, wheat, and rice**. We have isolated over three dozen border sequences (1 from Arabidopsis, 6 from potato, 17 from tomato, 2 from barrel medic, 1 from Alfalfa, 3 from barley, 2 from Brassica, 1 corn, 1 wheat, and 3 from rice). None of these sequences is identical to the Agrobacterium T-DNA border consensus. On the other hand, the P-DNA border-like sequences comprise a sequence that is highly homologous or identical to the core consensus sequence ANGATNTATN<sub>6</sub>GT. Please see Exhibit 5. The disclosed consensus sequences provide guidance concerning what constitutes a functional border-like sequence.

22. Our disclosed identification and functionality methodologies, and subsequent corresponding results, also informed us that an efficacious border from one plant species is very likely to be effective at transforming another plant species. For instance, we determined

Appl. No. 10/607,538  
Atty. Dkt. No. 058951-0167

that a pSIM100-OD-IPT-derived plasmid with the alfalfa border-like sequence CGGCAGGATGTATACAGAGGTATAC yielded transformation frequencies of about 127% in alfalfa and 168% in tobacco if compared to a conventional T-DNA vector control. The efficacy of that alfalfa border-like sequence is effective in the transformation of other plant species. Similarly, the tomato border-like sequence GGCCAGGATATATTTGTTGGTAATG supported a transformation frequency of about 100% in tobacco, tomato, and potato. Transformation methods for alfalfa and tomato were Agrobacterium-based, and referred to in paragraph 251. The Agrobacterium methods disclosed in the application specification were used to transform plants with the above alfalfa and tomato border-like sequence carried on modified pSIM 100-OD-IPT vectors. (See paragraphs 31, 341 and 344).

23. With respect to co-transfer, I appreciate that a co-transformation event that employs two T-DNA vectors can yield different transformation results compared to a T-DNA/P-DNA vector combination. Those different results, however, do not negate the predictability of those alternative co-transfer combinations. That is, even though a T-DNA/P-DNA co-transformation event may be "ten-fold" less effective than a T-DNA/T-DNA combination at integrating a desired polynucleotide (paragraph 385 and 389), the T-DNA/P-DNA arrangement still works. Indeed, we teach how to improve transformation efficiency by selectively enhancing desired integration events. See Examples 10-14.

24. It is clear, therefore, that the disclosed methods, the exemplified embodiments, and the extensive guidance provided by our specification for transforming plants with newly-identified border-like native polynucleotides, are predictable and successful. The method for identifying border-like sequences from plant genomes works. The method for testing the function of those border-like sequences works. The methods do not burden the skilled person with any undue experimentation. Not only do the methods work, and are easy to perform, but, importantly, they are predictable.

25. Accordingly, contrary to the Office's position, we do in fact provide the methods to identify and isolate a multitude of P-DNA border sequences. Furthermore, we also do in fact functionally evaluate those exemplified border-like sequences. Our methods are reproducible and predictably yield DNA fragments that can be readily tested for their functional ability to act as a border-like genetic element.

\* \* \*

26. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

**Declarant**

Full name of declarant: Caius Romprens

Declarant's signature



Date

8/29/06

Appl. No. 10/607,538  
Atty. Dkt. No. 058951-0167

Country of Citizenship: Netherlands

Residence Address: 3395 East Windsong, Boise ID 83712

## EXHIBIT 1

### DR. ROMMENS' CURRICULUM VITAE

#### Research Experience:

- Director Plant Sciences at J.R. Simplot Company, ongoing from September 2000.
  - Developed new methods for plant transformation and gene silencing (see, for instance, US Patent Applications 2003/0221213A1, 2005/0034188A1, and 2006/0156428A1)
  - Developed new strategies for crop improvement
- Team Leader Disease Control at Monsanto Company, from March 1995 until August 2000
  - Developed strategies for enhanced disease resistance in agriculturally-important crops (see, for instance, US Patents 6544733, 6506962, and 7030293)
  - Developed strategies for improved plant transformation (see, for instance, US Patent Application 2004/0133938A1).
- Postdoctoral Fellow at the University of California from January 1993 until February 1995.
  - Isolated various disease resistance genes (see US Patent 6245510)
- PhD Student at the Free University Amsterdam from Apr 1988 until December 1992.
  - PhD thesis of December 14<sup>th</sup>, 1992, entitled "Transposition of the maize Activator element in tomato"

#### Partial List of Publications:

**Rommens CM** (2006) Kanamycin resistance in plants: an unexpected trait controlled by a potentially multifaceted gene. *Trends Plant Sci* Jun 13; 11: 317-319.

Yan H, Chretien R, Ye J, **Rommens CM** (2006) New Construct Approaches for Efficient Gene Silencing in Plants. *Plant Physiol* Jun 9; 141: 1508-1518.

**Rommens CM**, Bougri O, Yan H, Humara JM, Owen J, Swords K, Ye J (2005) Plant-derived transfer DNAs. *Plant Physiol* 139: 1338-1349.

**Rommens CM** (2004) All-native DNA transformation: a new approach to plant genetic engineering. *Trends Plant Sci* 9 :457-464.

**Rommens CM**, Humara JM, Ye J, Yan H, Richael C, Zhang L, Perry R, Swords K (2004) Crop improvement through modification of the plant's own genome. *Plant Physiol* 135: 421-31.

**Rommens CM**, Kishore GM (2000) Exploiting the full potential of disease-resistance genes for agricultural use. *Curr Opin Biotechnol* 11: 120-125.

Gao AG, Hakimi SM, Mittanck CA, Wu Y, Woerner BM, Stark DM, Shah DM, Liang J, **Rommens CM** (2000) Fungal pathogen protection in potato by expression of a plant defensin peptide. *Nat Biotechnol* 18: 1307-1310.

Abad MS, Hakimi SM, Kaniewski WK, **Rommens CM**, Shulaev V, Lam E, Shah DM (1997) Characterization of acquired resistance in lesion-mimic transgenic potato expressing bacterio-opsin, *Mol Plant Microbe Interact* 10: 635-645.

Salmeron JM, Oldroyd GE, **Rommens CM**, Scofield SR, Kim HS, Lavelle DT, Dahlbeck D, Staskawicz BJ (1996) Tomato Prf is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the Pto kinase gene cluster. *Cell* 86: 123-133.

**Rommens CM**, Salmeron JM, Oldroyd GE, Staskawicz BJ (1995) Intergeneric transfer and functional expression of the tomato disease resistance gene Pto. *Plant Cell* 7: 1537-1544.

**Rommens CM**, Salmeron JM, Baulcombe DC, Staskawicz BJ (1995) Use of a gene expression system based on potato virus X to rapidly identify and characterize a tomato Pto homolog that controls fenthion sensitivity. *Plant Cell* 7: 249-57.

**Rommens CM**, van Haaren MJ, Nijkamp HJ, Hille J (1993) Differential repair of excision gaps generated by transposable elements of the 'Ac family'. *Bioessays* 15: 507-12.

**Rommens CM**, Munyikwa TR, Overduin B, Nijkamp HJ, Hille J (1993) Transposition pattern of a modified Ds element in tomato. *Plant Mol Biol* 21 1109-1119.

Rudenko GN, **Rommens CM**, Nijkamp HJ, Hille J (1993) Supported PCR: an efficient procedure to amplify sequences flanking a known DNA segment. *Plant Mol Biol* 21: 723-728.

**Rommens CM**, Rudenko GN, Dijkwel PP, van Haaren MJ, Ouwerkerk PB, Blok KM, Nijkamp HJ, Hille J (1992) Characterization of the Ac/Ds behaviour in transgenic tomato plants using plasmid rescue. *Plant Mol Biol* 20: 61-70.

**Rommens CM**, van Haaren MJ, Buchel AS, Mol JN, van Tunen AJ, Nijkamp HJ, Hille J (1992) Transactivation of Ds by Ac-transposase gene fusions in tobacco. *Mol Gen Genet* 231: 433-441.

**Rommens CM**, van der Biezen EA, Ouwerkerk PB, Nijkamp HJ, Hille J (1991) Ac-induced disruption of the double Ds structure in tomato. *Mol Gen Genet* 228: 453-458.

Haring MA, **Rommens CM**, Nijkamp HJ, Hille J (1991) The use of transgenic plants to understand transposition mechanisms and to develop transposon tagging strategies. *Plant Mol Biol* 16: 449-461. Review.

Haring MA, Gao J, Volbeda T, **Rommens CM**, Nijkamp HJ, Hille J (1989) A comparative study of Tam3 and Ac transposition in transgenic tobacco and petunia plants. *Plant Mol Biol* 13: 189-201.

**Exhibit 2**

“Plant-derived transfer DNA’s”, Plant Physiol. 139:1338-49, 2005.

# *Plant Physiology*

November 2005 • Volume 139 • Number 3



**Antifungal Proteins in the Nectar of Ornamental Tobacco**

An international journal devoted to basic research into how plants function, ranging from the molecular to the cellular to the whole plant levels, and including the interactions of plants with their biotic and abiotic environments.

## Editor-in-Chief

Donald R. Ort  
USDA/ARS University of Illinois, Urbana, IL, USA

## Associate Editors

Bonnie Bartel Rice University, Houston, TX, USA  
Herman Höfte INRA, Versailles, France  
Alan Jones University of North Carolina, Chapel Hill, NC, USA  
Maarten Koornneef Max Planck Institute, Cologne, Germany  
Robert L. Last Michigan State University, East Lansing, MI, USA  
Sheila McCormick USDA-ARS University of California, Berkeley, CA, USA  
Thomas Mitchell-Olds Max-Planck Institute for Chemical Ecology, Cologne, Germany  
John Ohlrogge Michigan State University, East Lansing, MI, USA  
Gary Stacey University of Missouri, Columbia, MO, USA  
Susanne von Caemmerer Australian National University, Canberra City, ACT, Australia

## Research Area and Sections

Signal Transduction and Hormone Action  
Cell Biology  
Development and Hormone Action  
Environmental Stress and Adaptation to Stress  
*Breakthrough Technologies and Genome Analysis*  
Systems Biology, Molecular Biology, and Gene Regulation  
Genetics, Genomics, and Molecular Evolution  
Biochemical Processes and Macromolecular Structures  
Plants Interacting with Other Organisms  
Bioenergetics and Photosynthesis  
Whole Plant and Ecophysiology

## Feature Editors

Richard Amasino University of Wisconsin, Madison, WI, USA  
Joseph J. Kieber University of North Carolina, Chapel Hill, NC, USA

Update and Meeting Report  
Update and Meeting Report

## International Board of Monitoring Editors

|  |   |   |  |
|--|---|---|--|
| Kozi Asada<br>Fukuyama, Japan                    | Edward Farmer<br>Lausanne, Switzerland      | Leon Kochian<br>Ithaca, NY, USA               | Pamela C. Ronald<br>Davis, CA, USA           |
| Antony Bacic<br>Parkville, VIC, Australia        | Peter Geigenberger<br>Golm, Germany         | Joachim Kopka<br>Golm, Germany                | David Salt<br>West Lafayette, IN, USA        |
| Julia Bailey-Serres<br>Riverside, CA, USA        | Thomas Girke<br>Riverside, CA, USA          | Ljerka Kunst<br>Vancouver, BC, Canada         | Daniel P. Schachtman<br>St. Louis, MO, USA   |
| Neil R. Baker<br>Essex, UK                       | Elzbieta Glaser<br>Stockholm, Sweden        | Norman G. Lewis<br>Pullman, WA, USA           | Danny J. Schnell<br>Amherst, MA, USA         |
| Philip W. Beraft<br>Ames, IA, USA                | Jeffrey Harper<br>Reno, NV, USA             | Hsou-min Li<br>Taipei, Taiwan                 | Paul Schulze-Lefert<br>Cologne, Germany      |
| Malcolm Bennett<br>Loughborough, UK              | Paul M. Hasegawa<br>West Lafayette, IN, USA | Chentao Lin<br>Los Angeles, CA, USA           | Edgar P. Spalding<br>Madison, WI, USA        |
| C. Robin Buell<br>Rockville, MD, USA             | Martha Hawes<br>Tucson, AZ, USA             | Elizabeth Lord<br>Riverside, CA, USA          | Dina St. Clair<br>Davis, CA, USA             |
| Edgar B. Cahoon<br>St. Louis, MO, USA            | Luis Herrera-Estrella<br>Irapuato, Mexico   | Francesco Loreto<br>Monterotondo, Italy       | Tuula Teeri<br>Stockholm, Sweden             |
| Judy Callis<br>Davis, CA, USA                    | N. Michele Holbrook<br>Cambridge, MA, USA   | M. David Marks<br>St. Paul, MN, USA           | Mary Tierney<br>Burlington, VT, USA          |
| Caren Chang<br>College Park, MD, USA             | Gregg Howe<br>East Lansing, MI, USA         | Susan Martino-Catt<br>Blacksburg, VA, USA     | Aart J.E. van Bel<br>Giessen, Germany        |
| Xuemei Chen<br>Riverside, CA, USA                | Dirk Inzé<br>Gent, Belgium                  | Ron Mittler<br>Reno, NV, USA                  | Carroll P. Vance<br>St. Paul, MN, USA        |
| Vitaly Citovsky<br>Stony Brook, NY, USA          | Georg Jander<br>Ithaca, NY, USA             | Mark O'Brian<br>Buffalo, NY, USA              | Andreas Weber<br>East Lansing, MI, USA       |
| Sacco C. de Vries<br>Wageningen, The Netherlands | Jacques Joyard<br>Grenoble, France          | Katherine Osteryoung<br>East Lansing, MI, USA | Mary Wildermuth<br>Berkeley, CA, USA         |
| Xinnian Dong<br>Durham, NC, USA                  | Elizabeth Kellogg<br>St. Louis, MO, USA     | Sara Patterson<br>Madison, WI, USA            | Shaul Yalovsky<br>Tel Aviv, Israel           |
| Natalia Dudareva<br>West Lafayette, IN, USA      | Anthony Kinney<br>Wilmington, DE, USA       | Hendrik Poorter<br>Utrecht, The Netherlands   | Kazuko Yamaguchi-Shinozaki<br>Ibaraki, Japan |
|  |   | Seung Rhee<br>Stanford, CA, USA               | Yunde Zhao<br>La Jolla, CA, USA              |

**Publisher**  
American Society of Plant Biologists,  
Crispin Taylor, Executive Director  
**Publications Director**  
Nancy A. Winchester; nancyw@aspb.org  
**Managing Editor**  
John Long; jlong@aspb.org

**Science Writer**  
Peter V. Minorsky; peminorsky@aspb.org  
**Production Manager**  
Jon Munn; jmun@aspb.org  
**Manuscript Manager**  
Leslie Ash Csikos; lcsikos@aspb.org



*Plant Physiology* (ISSN 0032-0889, online 1532-2548) is published monthly, in three volumes per year, by the American Society of Plant Biologists, 15501 Monona Drive, Rockville, MD 20855-2768, USA, and is printed by Dartmouth Journal Services, Orford, NH 03777. The institutional price for the print and online version is \$2160 per year and includes a subscription to both *The Plant Cell* and *Plant Physiology*; institutional online only pricing is \$1860; single copies may be purchased for \$75 each plus shipping. Members of the American Society of Plant Biologists may subscribe to *Plant Physiology* for \$175. Nonmember individuals may subscribe for \$325. For matters regarding individual subscriptions, contact Suzanne Cholwek, 15501 Monona Drive, Rockville, MD 20855-2768; telephone 301-251-0560, ext. 141, fax 301-251-6740, e-mail knoone@aspb.org. Notify ASPB within 1 month (domestic) or 2 months (overseas) of issue date, and defective copies or copies lost in the mail will be replaced. For matters regarding institutional subscriptions, contact ASPB Institutional Subscriptions, 15501 Monona Drive, Rockville, MD 20855-2768. Send all inquiries regarding display advertising to Kelly Taylor, Leonard Media Group, P.O. Box 220, Horsham, PA 19044; telephone 215-675-9133, ext. 226, fax 215-675-9376, e-mail kelly@leonardmedia.com. Periodicals postage paid at Rockville, MD, and at additional mailing offices. Postmaster: Send address changes to *Plant Physiology*, American Society of Plant Biologists, 15501 Monona Drive, Rockville, MD 20855-2768. The online version of *Plant Physiology* is available at [www.plantphysiol.org](http://www.plantphysiol.org). Permission to Reprint: Permission to make digital or hard copies of part or all of a work published in *Plant Physiology* is granted without fee for personal or classroom use provided that copies are not made or distributed for profit or commercial advantage and that copies bear the full citation and the following notice on the first page: "Copyright American Society of Plant Biologists." For all other kinds of copying, request permission in writing from Nancy A. Winchester, Publications Director, ASPB headquarters. © 2005 American Society of Plant Biologists. All rights reserved. Printed on acid-free paper. Printed in the USA.

## Plant-Derived Transfer DNAs

Caius M. Rommens\*, Oleg Bougri, Hua Yan, Jaime M. Humara, Joanna Owen, Kathy Swords, and Jingsong Ye

J.R. Simplot Company, Simplot Plant Sciences, Boise, Idaho 83706

The transfer of DNA from *Agrobacterium* to plant cell nuclei is initiated by a cleavage reaction within the 25-bp right border of Ti plasmids. In an effort to develop all-native DNA transformation vectors, 50 putative right border alternatives were identified in both plant expressed sequence tags and genomic DNA. Efficacy tests in a tobacco (*Nicotiana tabacum*) model system demonstrated that 14 of these elements displayed at least 50% of the activity of conventional *Agrobacterium* transfer DNA borders. Four of the most effective plant-derived right border alternatives were found to be associated with intron-exon junctions. Additional elements were embedded within introns, exons, untranslated trailers, and intergenic DNA. Based on the identification of a single right border alternative in *Arabidopsis* and three in rice (*Oryza sativa*), the occurrence of this motif was estimated at a frequency of at least  $0.8 \times 10^{-8}$ . Modification of plasmid DNA sequences flanking the alternative borders demonstrated that both upstream and downstream sequences play an important role in initiating DNA transfer. Optimal DNA transfer required the elements to be preceded by pyrimidine residues interspaced by AC-rich trinucleotides. Alteration of this organization lowered transformation frequencies by 46% to 93%. Despite their weaker resemblance with left borders, right border alternatives also functioned effectively in terminating DNA transfer, if both associated with an upstream A[C/T][C/G][A/T][T][G/T][C/T][G/T][C/G][A/C/T][C/T][A/T] domain and tightly linked cytosine clusters at their junctions with downstream DNA. New insights in border region requirements were used to construct an all-native alfalfa (*Medicago sativa*) transfer DNA vector that can be used for the production of intragenic plants.

*Agrobacterium*-mediated plant transformation is based on the transfer of single-stranded DNA from *Agrobacterium* plasmids to plant cell nuclei. DNA transfer is initiated by virD2-mediated cleavage within the 25-bp right border sequence (Wang et al., 1984; Scheffele et al., 1995). The *Agrobacterium* plasmids pTiC58 and pRi8196 contain additional sequences in the vicinity of the right border that enhance cleavage (van Haaren et al., 1987; Hansen et al., 1992). Upon initial cleavage, virD2 covalently binds to the 5' side, and the DNA unwinds toward the left border where a second cleavage reaction may occur. The released single-stranded DNA is coated with virE2 and processed for transfer via type IV-type secretion (Lessl and Lanka, 1994; Zupan and Zambryski, 1995). In the plant cell, nuclear localization signals of virD2 target the transfer DNA (T-DNA) to the nucleus for subsequent genomic integration (Shurvinton et al., 1992; Ziemienowicz et al., 2001). Because left borders often fail to function as sites for secondary cleavage, transformed plant cells frequently contain T-DNAs that are still attached to plasmid backbone sequences (Kononov et al., 1997). Furthermore, inadvertent right border activity of left borders can also result in the trans-

formation of plant cells that contain backbone DNA (Huang et al., 2004).

*Agrobacterium*-derived right border regions have been used extensively for the transformation of plants with foreign DNA. Resulting transgenic plants often displayed new and agronomically important traits such as herbicide and insect tolerance. However, the presence of foreign DNA in food crops is often perceived as undesirable, and the application of genetic engineering has therefore been limited to a small number of crops that are destined for feed, oil, fibers, and processed ingredients. In contrast, products closer to the table, such as fruits and vegetables, have been hindered in their transgenic development (Rommens, 2004). Public concerns were addressed through a proposal to diversify genetically modified crops based on the genetic distance between the source of new genetic material and target organism (Nielsen, 2003). According to this proposal, the introduction of foreign DNA creates "transgenic" plants, whereas "xenogenic" plants result from the insertion of laboratory-designed DNA for which no naturally evolved genetic counterpart can be found or expected. Some of these two groups of plants deviate substantially from what has been achieved through conventional breeding. In contrast, rearrangements of genomic material from within the same sexual compatibility group would create "intragenic" plants. Such modifications would often alter traits in a similar but more efficient and precise manner than that of conventional plant breeding.

One requirement for all-native DNA transformation is the replacement of *Agrobacterium* T-DNAs by plant-derived transfer DNA (P-DNA) sequences (Rommens, 2004). The concept of P-DNA-mediated

\* Corresponding author; e-mail crommens@simplot.com; fax 208-327-3212.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors ([www.plantphysiol.org](http://www.plantphysiol.org)) is: Caius M. Rommens (crommens@simplot.com).

Article, publication date, and citation information can be found at [www.plantphysiol.org/cgi/doi/10.1104/pp.105.068692](http://www.plantphysiol.org/cgi/doi/10.1104/pp.105.068692).

transformation has previously been demonstrated in potato (*Solanum tuberosum*). A 400-bp potato P-DNA delineated by regions that share homology with the left border of nopaline strains and the right border of octopine strains was effectively transferred from *Agrobacterium* to plant cells (Rommens et al., 2004). The potato P-DNA was subsequently used to introduce a silencing construct for a tuber-specific polyphenol oxidase gene into potato. Resulting intragenic plants displayed tolerance against black spot bruise sensitivity in impacted tubers.

Design of optimal P-DNA vectors requires a thorough understanding of the organization of *Agrobacterium* T-DNA border regions. Here, we isolated a variety of right border alternative sequences from plants, determined their original genomic context, and tested the efficacy of these elements in mediating DNA transfer. We also studied the role of flanking DNA, and identified sequences that support the initiation of DNA transfer. Furthermore, right border alternatives outperformed *Agrobacterium* left borders if placed between AT-rich DNA and cytosine clusters (C-clusters). Our new understanding of sequences involved in DNA transfer makes it possible to effectively identify and utilize P-DNAs.

## RESULTS

### Right Border Sequence Requirements

*Agrobacterium* right borders function in the initiation of DNA transfer. As shown in Figure 1A, these 25-bp sequences (Rb01–Rb07) share a highly conserved 13-bp 5'-ATATATCCTG-[C/T]-CA motif preceded by the more degenerate 12-bp consensus 5'-[A/C/G]-[A/T]-[A/T]-[G/T]-AC-[A/C/T]-N-[C/G/T]-[A/C/G]-[A/C/G]-N (CON1). The sequence plasticity of right border alternatives was studied by testing synthetic elements containing single mismatches to CON1 (Sy01–Sy13) for their ability to mediate plant transformation. The elements, flanked by the 109-bp *Agrobacterium* pTi15955 sequence upstream from Rb02, were inserted into a plasmid containing an expression cassette for the neomycin phosphotransferase (*nptII*) selectable marker gene. *Agrobacterium* strains carrying the resulting test plasmids were used to infect tobacco (*Nicotiana tabacum*) explants. Two weeks after infection, the average numbers of calli per explant were compared to those produced with control plasmids containing Rb01 or Ct01, whereby  $15.3 \pm 0.5$  calli/explant = 100% transformation efficiency. As shown in Figure 1B, all synthetic right border alternatives had maintained the ability to mediate DNA transfer. However, base substitutions C6A, A13C, C19G, C20G, and T21A of Sy03, Sy07, Sy11, Sy12, and Sy13, respectively, lowered transformation frequencies more than 5-fold, implying important negative effects of these specific modifications on the ability of right border alternatives to initiate DNA transfer.

Right border alternative sequence requirements were further determined by testing the efficacy of plant sequences that either match or resemble CON1 (Fig. 1C). In addition to the previously characterized element of the potato P-DNA (Rommens et al., 2004), designated here as St01, a large number of new putative right border alternatives was either amplified from genomes or identified in publicly available DNA databases (see "Materials and Methods"). Among these sequences, only the *Arabidopsis* At01 fully matched the *Agrobacterium* right border consensus. However, this element displayed a lower activity (65%) than the control borders, suggesting that the guanine base at position +4 (G4) of At01 and poorly characterized right borders such as Rb03 (Jouanin et al., 1986) is not as effective as T4 of the more commonly used Rb01 and Rb02.

Despite the presence of one to three mismatches, 13 right border alternatives from potato (St02, St03, and St04), tomato (*Lycopersicon esculentum*; Le01, Le02, and Le03), pepper (*Capsicum annuum*; Ca01), alfalfa (*Medicago sativa*; Ms01), barley (*Hordeum vulgare*; Hv01), rice (*Oryza sativa*; Os01, Os02, and Os03), and wheat (*Triticum aestivum*; Ta01) displayed between 50% and more than 100% activity compared to the Rb01 and Ct01 control sequences. The functionality of these novel sequences indicates that *Agrobacterium* did not exploit the full potential of border sequence variation. In contrast to the highly effective right border alternatives, 36 additional elements from a variety of plant species contained multiple mismatches and/or point deletions, and either displayed low activities (<50%) or appeared nonfunctional. Mismatches that reduced transformation frequencies most dramatically include, apart from those mentioned above, A5G and C6G. By comparing the sequences of the two different groups of elements, the consensus for plant-derived right border alternatives (CON2) was determined as 5'-[A/C/G]-[A/C/T]-[A/C/T]-[G/T]-ACNNNNNNA-[G/T]-A-[A/C/T]-[A/G]-TCCTG-[C/G/T]-[A/C/G]-N (Fig. 1D).

### Plant-Derived Right Border Alternatives Often Represent Sequences Involved in RNA Processing

To determine the possible function of right border alternatives in plants, we analyzed the genomic context of some of the most effective sequences. Interestingly, several expressed sequence tag (EST)-derived elements were found to be associated with intron-exon junctions (Fig. 2A). For example, genomic DNA corresponding to the inverse complement of tomato Le02 contains an intron positioned between the bases AG and GT, which represent the dinucleotides that most frequently flank plant introns (Long and Rosenberg, 2000). Associations with intron-exon transitions were also demonstrated for tomato Le04 and rice Os01. In the case of Le04, the first six nucleotides 5'-TCAGAG are located in an intron (with acceptor site underlined), and the subsequent 19 nucleotides 5'-GATATATTAGACAGTT-TCC are part of the downstream exon of a gene with

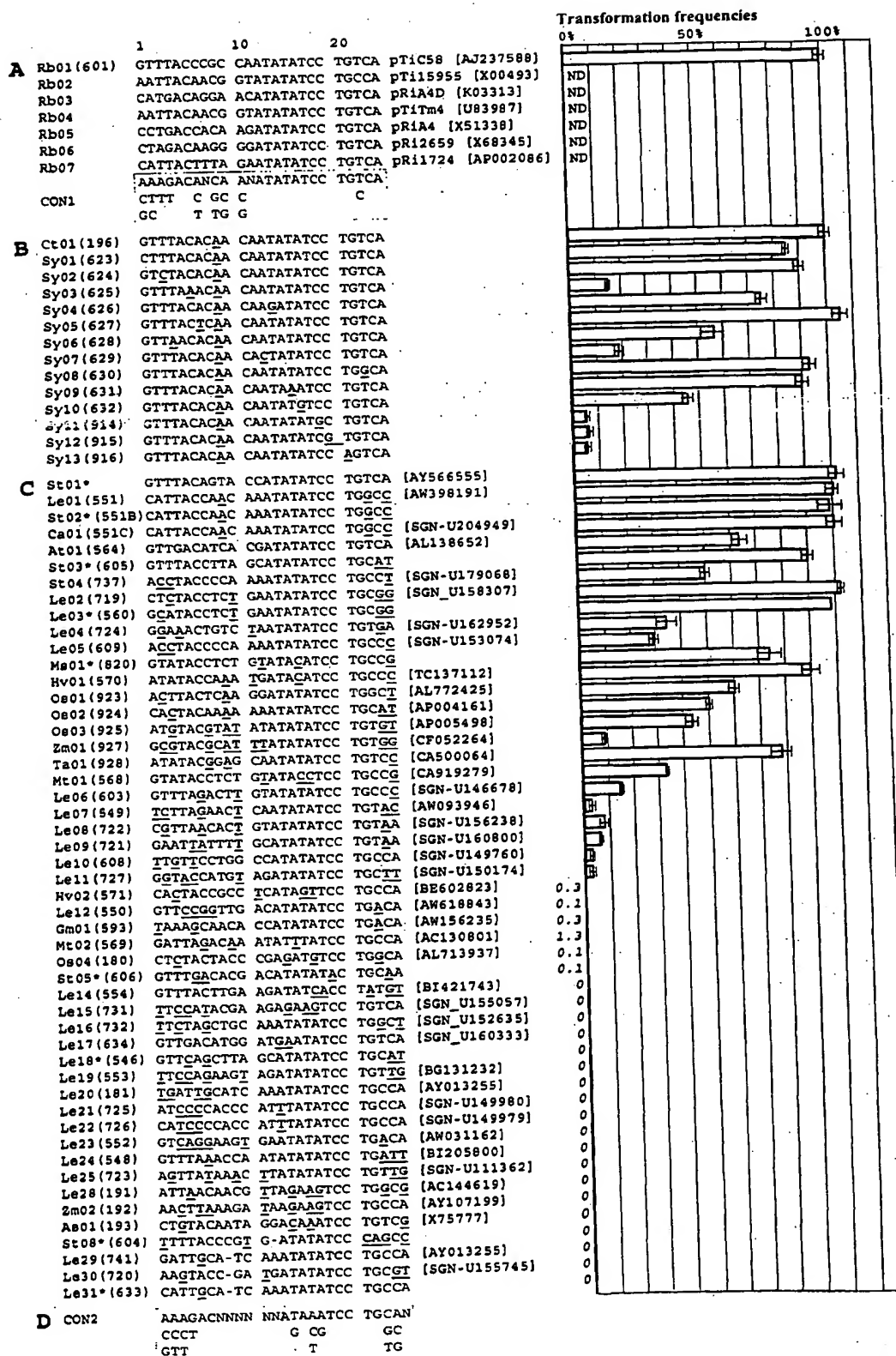


Figure 1. Right border sequence requirements. Elements are shown in blue with mismatches to the consensus of *Agrobacterium* right borders (CON1) indicated in red. Corresponding accession numbers, if available, are shown between brackets, and associated pS1M vector numbers are indicated between parentheses. A, Right borders from *Agrobacterium tumefaciens* (Rb01

unknown function. Similarly, Os01 defines the transition from exon (5'-AGCCAGGATATATCCTTGA) to intron (5'-GTAAGT, with donor site underlined) of the rice cleavage stimulation factor 77 gene. Another EST-identified right border alternative that is associated with an intron is Le01. This element is located immediately downstream from the intron acceptor site of the tomato phosphatase 2C gene. Identical elements from potato (St02) and pepper (Ca01) were determined to represent similar junctions (data not shown). Interestingly, all these six right border alternatives were positioned in the same inverse complementary position relative to the orientation of the gene. The generally GT-rich alternative strands contain, apart from AG and GT splice sites, binding sites for Ser/Arg-rich splicing enhancers 5'-[C/G]-[A/G]-[C/G]-A-[C/G]-G-[A/T] (Liu et al., 1998; Reddy, 2004).

In contrast to the above-described first group of right border alternatives, the position and/or structure of a number of other elements did not suggest an immediate link with transcript processing. As depicted in Figure 2B, three sequences were derived from untranslated trailers (3'UTRs). The barley Hv01 element is located within the 3' UTR of the *LepA* gene, and the *Medicago truncatula* Mt01 and maize (*Zea mays*) Zm01 are positioned in untranslated sequences of the *AIP* gene and a putative 8-amino-7-oxononanoate synthase gene, respectively. Furthermore, the Arabidopsis At01, rice Os02, and wheat Ta01 elements are all embedded within exon sequences of genes. In the case of rice Os03, the element is fully embedded within the 4.3-kb 11th intron of a rice *Sec8*-like gene. And finally, some right border alternatives that had been PCR amplified from plant genomes represented DNAs lacking an apparent association with genes. Like the previously isolated St01 (Rommens et al., 2004), potato St03 appears positioned in the AT-rich intergenic DNA region (Fig. 2B). Collectively, our results demonstrate that plant-derived right border alternatives can be derived from a variety of genetic elements. One group of elements is predominantly associated, in the inverse complementary (antisense) orientation, with the transcribed strands of genes. In agreement with the transcription direction rule (Szybalski et al., 1966), these anti-sense borders are poor in guanine content.

#### The Influence of Flanking DNA on the Efficacy of Right Border Alternatives

Test plasmid pSIM551 contained St02 linked to T-DNA-flanking sequences of *Agrobacterium* pTi15955.

This arrangement placed the potato right border alternative at a distance of 12 bp from the overdrive, an element that was reported to promote DNA transfer (van Haaren et al., 1987). Although the overdrive element is believed to function in a position-independent manner with respect to the right border (Shurvinton and Ream, 1991), we found that a single basepair insertion between St02 and upstream DNA reduced transformation frequencies of pSIM579 about 2-fold (Fig. 3A). Furthermore, the 5'-CAA trinucleotide insertion of pSIM578 had an even greater negative effect on the efficacy of transformation, lowering it to 35%.

To study the molecular basis of the apparent overdrive-St02 spacing requirement, we compared the pSIM551 sequence with corresponding T-DNA flanking regions of *Agrobacterium* plasmids. The aligned 47-nucleotide sequences generally contained cytosine or thymine residues at conserved four-nucleotide intervals, separated by adenine-rich (46%) trinucleotide segments (Fig. 3A). This arrangement resulted in a high occurrence of AC dinucleotide repeats (27%), which approached that of the overdrive element itself (42%). Whereas the upstream AC-rich (ACR) domains from both *Agrobacterium* plasmids and pSIM551 comprised at least six pyrimidine residues at conserved positions, the impaired activity of pSIM578 and 579 was correlated with sequences that contained four and five such residues, respectively (Fig. 3A). Additional evidence for the importance of correctly spaced pyrimidines was obtained by analyzing pSIM580, which contained the pentanucleotide 5'-ACCAA insertion between St02 and upstream DNA. Maintenance of six pyrimidines at conserved positions in this plasmid was associated with the same DNA transfer activity as that of the original vector pSIM551 (Fig. 3A).

To further test the functional significance of correctly spaced pyrimidines, the pTi15955-derived sequence of pSIM551 was replaced by the corresponding ACR region of *Agrobacterium* pRi2659 (Hansen et al., 1992). Immediate linkage with St02 in pSIM844 resulted in high transformation frequencies (126%; Fig. 3A). However, disruption of the pyrimidine spacing by a single basepair insertion (in pSIM827) almost fully eliminated the activity of the right border alternative, reducing it down to 7%.

Having correlated the original spacing of pyrimidines with efficient DNA transfer, we now also tested the functional relevance of adenine-rich trinucleotides. For this purpose, the ACR domain of pSIM551 was replaced with a tomato DNA fragment carrying nine pyrimidines at conserved positions but lacking a high

Figure 1. (Continued.) and Rbo2), *Agrobacterium rhizogenes* (Rb03, Rbo4, Rb05, Rb06, and Rb07), and *Agrobacterium vitis* (Rb04). B, Synthetic elements. Ct, Synthetic control border; Sy, synthetic borders containing a single mismatch to CON1. Horizontal bars show transformation frequencies compared to those supported by Rb01. C, Plant-derived right border alternatives. Sequences that were isolated by employing PCR/inverse PCR approaches are indicated with asterisks. D, Overall right border/right border alternative consensus (CON2). Transformation frequencies are shown as percentages of that of the Rb01 control vector and represent the mean  $\pm$  SE of at least three experiments.

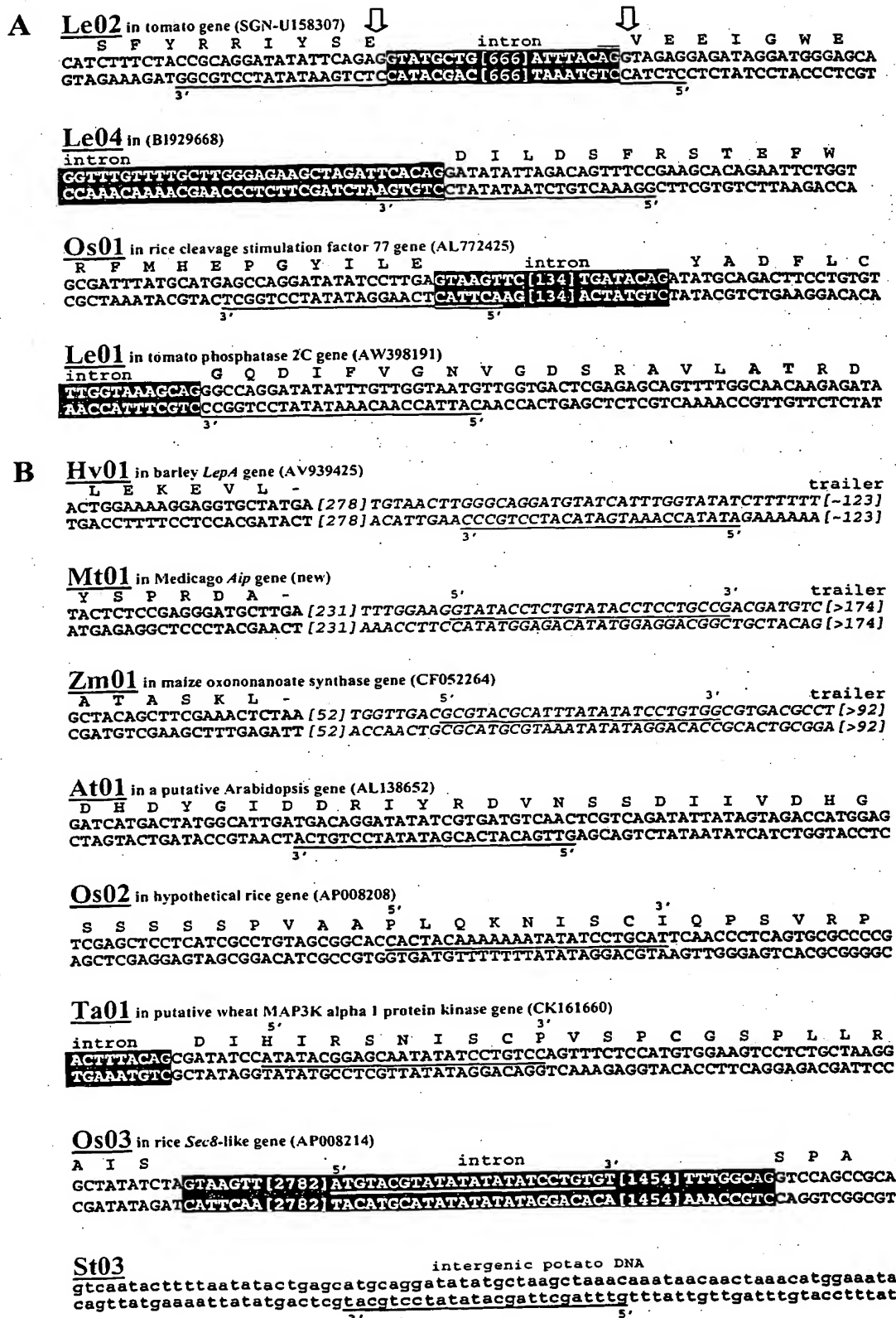


Figure 2. Genomic context of plant-derived right border alternatives. A, Elements (underlined) at intron-exon junctions. Arrows indicate the most frequent intron-flanking dinucleotides. B, Right border alternatives positioned within untranslated leaders (depicted in gray italics), exons, introns, and intergenic DNA (lowercase). Exon sequences are shown in black with amino acids that they encode in red. Introns are boxed. The number of basepairs omitted from the text is shown between brackets. Accession numbers of genomic clones, if available, are shown between parentheses.

**A**

5' -4 +1 Right Border/RBA

pSIM551 CTTAGAGATCTCAAAACAAACATACACAGCGACTTATTCACTAGTCATTACCAACAAATATATCCTGGCC (100 ± 5)

pSIM579 TTAGAGATCTCAAAACAAACATACACAGCGACTTATTCACTAGTCATTACCAACAAATATATCCTGGCC (48 ± 1)

pSIM578 AGAGATCTCAAAACAAACATACACAGCGACTTATTCACTAGTCATTACCAACAAATATATCCTGGCC (35 ± 4)

pTi15955 AGAAACAATCAAAACAAACATACACAGCGACTTATTCACTAGTCATTACCAACAAATATATCCTGGCC (ND)

pTiC58 GCCCTTTTAAATATCCGATTATTCTAATAACGCTCTTTCTCTTAGGTTTACCCGCCAATATATCCTGTCA (ND)

pRi2659 TGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAAGGGATATATCCTGTCA (ND)

pRiA4 TAACAATTGAACAATTGAACAATTGAACAATTGAACAATTGAACAACAATGACAGGAACATATATCCTGTCA (ND)

pRi8196 TAGACATTGCACATCCAAAGGCAGGCACGTACAAACGAATTTATTTAGCCGACAACGGAATATATCCTGTCA (ND)

pRi1724 GAAGGCACGAAGGCACGAAGGCACGAAGGCACGAAGGCACGAAGGCACATTACTTTAGAATATATCCTGTCA (ND)

pRiDB18 TCATCACC~~CGCTCCT~~AAACAAACATACCTCCACACAATTTATCTACCTGACCACAAGATATATCCTGTCA (ND)

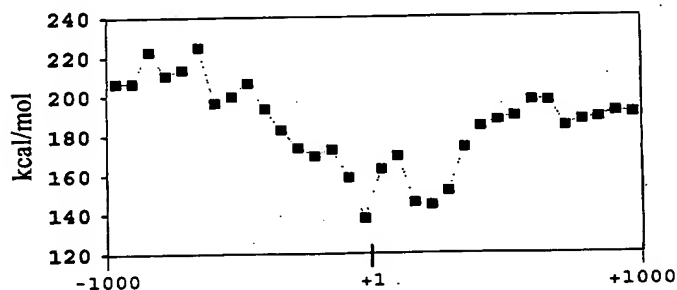
pSIM580 AGATCTCAAAACAAACATACACAGCGACTTATTCACTAGTCATTACCAACAAATATATCCTGGCC (120 ± 5)

pSIM844 TGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACAAATATATCCTGGCC (126 ± 5)

pSIM827 CTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACTGACGAACAAATATATCCTGGCC (7 ± 2)

pSIM581 TGTCTTTATCTCTTGTGTCAAAACACTGCTCTCGAGTCGAGTCACCAACATTACCAACAAATATATCCTGGCC (22 ± 2)

ACR domain

**B****C**

OD CAAACAAACATACACAGCGACTTA

+26

pTiC58 GTCAGCATCATCACACCAAAAGTTAGGCCGAATAGTTTGAAATTAGAAAGCTCGCAATTGAG (100 ± 4)

pTi15955 AACACTGATAGTTTAAACCGAAGGCGGGAACGACAATCTGATCATGAGCGGAGAATTAAGGG (ND)

pRi2659 AATAACAATCTCATGTTAGGTAATAATATCACCCAATCAACGCGGCCACGCAATTAACCTGGT (ND)

pRiA4 GTCATCAGCAACAGCAACAGTAGTTATTGTCTGTGAAGATATGTAGGTACCTTTCACCCAC (ND)

pRi8196 GCACTAATATAAGAAATGTCCTGTGACACTAATATAAGAAATGTCCTGCTTTGACGCAAGTG (ND)

pRi1724 AACCTATTCGTTAATAGGACGTCGTACCTACTTCCCTTCCAGCGCAGCAACAGTAGGTCTG (ND)

pSIM108 GAGGTATAGAGGCATGACTGGCATGATCACTAAATTGATGCCACAGAGGAGACTTATAACCT (112 ± 5)

pSIM551 GGGCCCCGTACCCGGGGATCAATTCCCGATCTAGTAACATAGATGACACCGCGCGGATAATT (100 ± 5)

pSIM920 GGGCCCCGTTCCTGGGGATCAATTGGGCCCCGTACCCGGGGATCAATTCCCGATCTAGTAACA (60 ± 4)

pSIM781 GGGCCCCGTACCCGGAGGAGACTCCGATCTACGGCGCCAAATTCAAGGACGGAGAACTTCGTG (89 ± 4)

pSIM582 CTGAGGACATTGAGAAGATTGGTTATATCCTCTTTCAAGACGCTAAGCAATTAGTGATGCAAA (59 ± 4)

pSIM793 GAGGTATAGAGGCATGTCCTGGCGTGATCACTAAATTGATGCCCGCAGAGGGGACTTATAACAT (105 ± 4)

pSIM843 GGGCCCCGTACCCGTTAGGGCTAGCCCGAAAGGGCCGCGGCAGCCCGTTAGCCCGCATAAC (168 ± 12)

Figure 3. Sequences flanking right border alternatives. A, Upstream sequences display a conserved organization of cytosine/thymine residues separated by adenine-rich trinucleotide spacers. The overdrive sequence of pTi15955 is underlined (dotted).

percentage of adenine residues in the intervals. The resulting vector pSIM581 displayed only 22% of the transformation efficacy of pSIM551, indicating that adenine-rich intervals and/or AC repeats play a role in the functional activity of the ACR domain (Fig. 3A).

Because adenine-rich DNA is often associated with low helical stability regions, we determined the helical stability profile of pSIM551 using WEB-THERMODYN (Huang and Kowalski, 2003). This analysis identified a 120-bp sequence immediately upstream from the St02 cleavage site, and including the ACR domain, to represent the lowest helical stability region of the pSIM551 backbone (Fig. 3B; data not shown). The association of an easily unwound DNA region immediately upstream from the right border alternative may be functionally relevant because *Agrobacterium* Ti and Ri plasmids contain similar low helical stability regions at their right borders. For instance, pTiC58 contains a 120-bp region preceding the border with a stability of 116 kcal/mol. Analogous to the association of low helical stability regions with the initiation of plasmid replication (Natale et al., 1993), these upstream DNAs may be involved in the initiation of DNA transfer.

Given that upstream DNA sequences adjacent to St02 influenced transformation efficacy, we sought to test the effect of downstream modifications. As shown in Figure 3C, sequence analyses identified decamers that shared the consensus 5'-[A/C/T]-[A/C]-[A/C/T]-[A/G/T]-[A/T]-T-[A/C]-G-[G/T]-[G/T] with the 5' part of the overdrive, and were positioned at a distance of one to 27 nucleotides from the right border. This "downstream from right border" (DR) domain was also identified in both the potato-derived T-DNA (Rommens et al., 2004) of pSIM108 and test vectors such as pSIM551 (Fig. 3C). An increase in the spacing between Le01 and DR domain from 24 nucleotides in pSIM551 to 48 nucleotides in pSIM920 lowered transformation frequencies by 40% (Fig. 3C), indicating that the supporting function of the DR domain on border activity is spacing dependent.

Because downstream DNA sequences represent the actual T-DNA that is intended for plant transformation, we replaced the original bacterial sequences of pSIM551 with two unique potato DNA fragments. The pSIM551-derivative pSIM793, which contained a DR domain at 27 nucleotides from Le01, yielded about the same transformation frequency as pSIM551. In contrast, the potato DNA fragment of pSIM582, which contained a DR domain with several mismatches to

the consensus, displayed only 59% activity. Interestingly, replacement of Le01-flanking DNA sequences by an alfalfa DNA fragment that contained two different DR domains triggered unusually high transformation frequencies for the resulting vector pSIM843 (168%; Fig. 3C). We conclude that sequences flanking right border alternatives play an important role in supporting plant DNA transfer. These sequences comprise upstream ACR and downstream DR domains.

#### Substitution of Left Borders by Right Border Alternatives

The above-described studies had shown that CON2-matching 25-bp elements function as effective right border alternatives if flanked by sequences that support their activity. Would it also be possible to use such elements as left border alternatives? The answer to this question was not immediately obvious because the intended function of left borders, namely secondary cleavage of the T-DNA, is different from that of right borders, which mediate primary cleavage reactions. As shown in Figure 4A, this functional difference is associated with a divergent sequence organization at and around the border sites. In contrast to right borders, left borders (1) are preceded by AT-rich DNAs comprising an "upstream from left border" (UL) domain on either DNA strand with the consensus sequence A[C/T]T[C/G]A[A/T]T[G/T][C/T][G/T][C/G]A[C/T][C/T][A/T], (2) share a more conserved consensus sequence, 5'-[A/G]TTTACA[A/C/T][A/C/T][A/C/T][C/G]AATATATCCTGCC[A/G], and (3) are linked to downstream plasmid backbone DNA by C-clusters that conform to the consensus CCN<sub>1-11</sub>CCN<sub>1-11</sub>CCN<sub>1-11</sub>CC.

Direct evidence for the role of the C-cluster organization in supporting left border activity was obtained by comparing the fidelity of DNA transfer for pSIM831 and 829. Both vectors contained an expression cassette for the *nptII* gene preceded by DNA regions comprising St02 as right border alternative, and were confirmed to support the same high transformation frequencies as pSIM551 (data not shown). The vectors also contained almost identical DNA regions for secondary cleavage, which differed only in that pSIM829 contained a 10-bp insertion in the fourth left border-associated C-cluster (Fig. 4B). The effect of this small change was assessed by classifying regenerated shoots in three groups based on PCR analyses. The first "T" group only contained the intended T-DNA, and would therefore be predicted

Figure 3. (Continued.)

Direct repeats are indicated with gray arrows. Transformation efficacies are shown between parentheses as percentages of controls, and represent the mean  $\pm$  SE of three experiments. "+1" indicates the position of the first base of the right border or right border alternative. ND, Not determined. B, Helical stability profile (kcal/mol) across the extended 2-kb St02 region of pSIM551 with 60-bp step size and 120-bp window size. C, Downstream sequences comprise a DR domain (red with mismatches in black). OD, Overdrive. Plasmids pSIM781, 793, and 843 contain DNA fragments from a potato homolog of AY566555, a potato homolog of AY972080, and an alfalfa homolog of *M. truncatula* AC131026, respectively. Plasmid pSIM582 contains Le01 flanked by the same tomato DNA sequence that flanks the element in its original genomic context. Transformation frequencies are indicated between parentheses as percentages of controls and represent the mean  $\pm$  SE of three experiments.

**A**

pTi15955

TCTCCATATTGACCATCATACTCATTGCTGATCCATGTAGATTTCCCGGACATGAAGCC  
 ATTTACAATTGAATATATCCTGCGCGCTGCGCTTTGCACCC

pTiC58

TGAATTCAGTACATTAAAAACGTCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTT  
 GTTTACAACAATATATCCTGCGACGAGCGAACAGCTCCCCGACC

pRi2659

ATCTGGTAATATAGCAAAAACGTGCTCAAAAATCGCTTCAAAGCTCTTGTACTTAGCTC  
 GTTTACAACAATATATCCTGCGACCCC

pRiA4

TACATTTTATATTGATAAAGCATGCGTTAAAACGACTTCGCATGTCCATATCTAATCT  
 GTTTACATCACAATATATCCTGCGACCAAGGAGCGACGCCCTTCTGGCC

**B**

pSIM831

AAATCTGATTGATAAAGGATCGATCCTCTAGAGTCGACCTGCAGTACTTACGTACAATT  
 GTTTACAACAATATATCCTGCCACCGGATATATTGCTAGGAGCCAGCAACAGCT  
 CCCCAGACC

pSIM829

AAATCTGATTGATAAAGGATCGATCCTCTAGAGTCGACCTGCAGTACTTACGTACAATT  
 GTTTACAACAATATATCCTGCCACCCCTAGGAGCCAGCCAACAGCTCCCCGACC

**C**

pSIM108

TCCTTCATAGCTACACTTTCTAAAGGTACGATAGATTTTGGATCAACCACACACTT  
 CTTTACAACGGTATATATCCTGCCAAAGCTTCCAGCCAGCCAACAGCTCCCCGACC

pSIM843B

GTAAAAAATAAAGTGAAATTCATGAATTAACACAAATATAAATGTAATATAAAATT  
 GTATACTCTGTATACATCCTGCCGCAAGCTTCCAGCCAGCTAGGAGCCAGCCAACA  
 GCTCCCCGACC

pSIM849

AATGGAGGTAAGTGTTTCTGCTCAGTGCTGATAGATGTAATATCTCTGTTATGAAGCC  
 GTATACTCTGTATACATCCTGCCGGGATGTATACCTTAGGCCAGCCAACAGCTCCC  
 CGACC

pSIM781

TGTTGAAGGCTTGATGTGATTAAGAAGGCCGAGGCTGTTGGATCTAGTTCTTGAAGTT  
 CATTACCAACAATATATCCTGCCCCCCCTAGGAGCCAGCCAACAGCTCCCCGACC

## Transformation-%

|          | T      | TB     | B      |
|----------|--------|--------|--------|
| pTi15955 | 26 ± 8 | 55 ± 9 | 19 ± 6 |
| pTiC58   | ND     |        |        |
| pRi2659  | ND     |        |        |
| pRiA4    | ND     |        |        |
| pSIM831  | 41 ± 3 | 52 ± 5 | 7 ± 2  |
| pSIM829  | 17 ± 2 | 68 ± 5 | 15 ± 4 |
| pSIM108  | 41 ± 5 | 48 ± 6 | 11 ± 3 |
| pSIM843B | 34 ± 4 | 53 ± 6 | 13 ± 2 |
| pSIM849  | 10 ± 2 | 83 ± 7 | 7 ± 2  |
| pSIM781  | 11 ± 1 | 58 ± 3 | 31 ± 2 |

Figure 4. DNA sequences flanking left borders and left border alternatives. A, Left border region of Ti and Ri plasmids. B, Modified left border regions. C, Regions comprising St01, Ms01, or St02. Upstream DNA is shown in blue with UL domains indicated in bold. Left borders and left border alternatives are highlighted in gray. C-clusters are boxed. Frequencies of transgenic plants containing the designated T-DNA delineated by borders or border alternatives ("T"), the T-DNA still attached to backbone sequences ("TB"), and backbone only ("B") are shown on the right and represent the mean ± SE of three experiments. ND, Not determined.

to have arisen from primary cleavage events at the right border followed by secondary cleavage at the left border. Plants containing both the T-DNA and additional backbone DNA sequences were classified in a second "TB" group, and most likely represented events where the second copy of the border alternative failed to function in terminating DNA

transfer. The third "B" group of events only contained backbone DNA and probably arose from initial cleavage reactions at the second St02 copy. This genotype classification demonstrated that pSIM831 was more than twice as effective as pSIM829 (41% versus 17%) in producing "T" events (Fig. 4B).

Efficacy of right border alternatives as sites for secondary cleavage was studied by testing pSIM108 and 843B. The vectors contained St01 and Ms01, respectively, as right border alternative. The downstream region of pSIM108, shown in Figure 4C, contained a second copy of St01 inserted between (1) AT-rich (62%) DNA derived from the terminator of the potato ubiquitin-3 gene (Garbarino and Belknap, 1994) containing a UL domain, and (2) plasmid backbone DNA comprising five C-clusters. Similarly, the DNA region intended for secondary cleavage in pSIM843B contained a second copy of Ms01 preceded by an AT-rich (87%) alfalfa DNA fragment and followed by downstream C-clusters (Fig. 4C). The binary vector pSIM401, which contained the extended left border region of pTiC58, was used as control. PCR genotyping demonstrated that both pSIM108 and 843B yielded even higher frequency of backbone-free transformation events (41.1% and 33.9%) than obtained with the control (26.0%), thus indicating that right border alternatives can be used to replace left borders. However, subsequent experiments also showed that such replacements need to be carried out cautiously. A modification of pSIM843B that both eliminated the UL domain and altered the spacing of C-clusters dramatically lowered the frequency of desired "T" transformation events for the resulting vector pSIM849 to 10.2% (Fig. 4C). This reduced frequency was associated with an about 2-fold increased transfer of DNAs that are still attached to their vector backbones, indicating that the modifications of flanking DNA interfered with effective secondary cleavage at the second Ms01 copy. Similar alterations of DNA surrounding the second St01 copy of pSIM108 resulted in an almost 4-fold reduced transformation efficacy of the resulting vector pSIM781 (Fig. 4C).

Collectively, our data demonstrate that right border alternatives can be used to replace left borders if associated with the upstream UL domain and downstream C-clusters. Even small changes in this organization were found to have a profound effect on the frequency of backbone-free plant transformation. Replacement of the internal *nptII* gene expression cassette of pSIM843B by alfalfa DNA would make it possible to produce intragenic alfalfa plants.

## DISCUSSION

This article describes 14 new and functionally active right border alternatives from dicotyledonous (potato, tomato, pepper, alfalfa, and Arabidopsis) and monocotyledonous (rice, barley, and wheat) plant species. Because analyses of the 125-Mb haploid Arabidopsis genome and 3-fold larger rice genome uncovered a single Arabidopsis element and three elements in rice, the occurrence of right border alternatives was estimated at a frequency of at least  $0.8 \times 10^{-8}$ . Thus, the 8-fold larger tomato and potato genomes are expected to each contain more than eight such se-

quences, several of which are described in this article. Continued plant genome sequencing and data mining efforts are anticipated to result in the identification of many additional right border alternatives from these and other plants. Alternatively, it is possible to link the sequences of at least two different plant ESTs together and create new synthetic right border-like elements. We have tested two such elements, 5'-GGACAGGATATATAA AGTGTAACAAAC and 5'-GGACAGGATATATAAGTGTAACAAAC, the sequences of which were recently disclosed ([www.crop.cri.nz](http://www.crop.cri.nz)). In our tobacco transformation system, these elements supported 86% and 7%, respectively, of the transformation efficacy of conventional right borders (C.M. Rommens, O. Bougri, and J.M. Humara, unpublished data).

In addition to plant-derived right border alternatives, such elements were also identified in fungi and mammals. For instance, mouse accession number AC110541 contains the sequence 5'-AGGCAGGACTAATGTGGTGTAAC. With the discovery that *Agrobacterium* can be used for transformation of a wide variety of species (Kunik et al., 2001; Casas-Flores et al., 2004), the new right border alternatives may be used to extend the concept of all-native DNA transformation (Rommens, 2004) to eukaryotes other than plants.

Many of the plant-derived right border alternatives described in this article are associated with transcribed DNA. These sequences represent untranslated trailers, introns, exons, as well as intron-exon junctions. The latter group of elements was identified most frequently and suggests a function of the GT-rich antisense border sequences in intron splicing. Because guanines and thymines each have a choice of two bases to pair with, guanines with cytosines and thymines, and thymines with adenines and guanines, the antisense borders are ideal for secondary structure formation, which is generally believed to play an important role in pre-mRNA splicing (Patterson et al., 2002). Interestingly, we found the inner part of the Le01 intron of the phosphatase 2C gene to share homology with introns of many other genes, such as a tomato permease gene, a Suc synthase gene, a TPR domain-containing protein gene, and a receptor kinase-like protein gene from potato (C.M. Rommens, unpublished data). Although it is conceivable that such introns evolved from ancient insertions of the same mobile intron (Long and Rosenberg, 2000), alignment of corresponding intron-exon junctions did not reveal any conserved sequences. Thus, Le01 does not appear to represent the remnants of an ancient proto-splice site. Sequences that flank the right border alternatives in their original plant genomic context lacked homology with known bacterial DNA (C.M. Rommens, unpublished data), indicating that these regions also did not originate from early bacterium-mediated transformation events.

Efficacy tests of border-like sequences made it possible to define a consensus sequence for right border alternatives. This consensus allows considerable degeneracy, especially in the 5' terminus, and confirms the previous notion that mismatches with

*Agrobacterium* right border sequences are often tolerated (Rommens et al., 2004). It also suggests that both right borders and right border alternatives contain two separate domains, a possible downstream binding site for virD2 preceded by a domain that may function as site for an accessory protein. Because the latter domain resembles the relaxed consensus binding site of the eubacterial initiator protein of DNA replication (DnaA), 5'-[C/T]-[C/T]-[A/C/T]-T-[A/C]-[C/T]-A-[A/C]-A (Roth and Messer, 1998; Speck and Messer, 2001), this or a related protein may be involved in supporting local DNA unwinding.

The sequence immediately upstream from right borders and right border alternatives represents a low helical stability region that may be involved in DNA unwinding as well. In *Agrobacterium* plasmid pTi15955, this region contains a 25-bp sequence (overdrive) that has been reported to promote right border activity (van Haaren et al., 1987). Our work demonstrates that the overdrive is part of a larger ACR domain that is conserved among *Agrobacterium* Ti and Ri plasmids. We found the enhancing activity of this region to require at least six pyrimidine residues at conserved positions. The sequence and spacing requirements of the ACR domain may explain why certain previously characterized Ti plasmid mutants displayed reduced virulence (Shurvinton and Ream, 1991). For instance, the impaired activity of mutant WR1803 is associated with an insufficient number of pyrimidine residues at conserved positions, and the low efficiency of WR1804 is linked to a lack of adenine-rich spacers. This weakly conserved repetitive and spacing-sensitive character of the ACR domain shares a similar architecture with the repeat region that is linked to the replication origin of, for instance, repABC-type plasmids. Like these origin-associated repeats, the ACR domain could play a similar role in DNA unwinding (Jakimowicz et al., 2002; Betteridge et al., 2004; this work). A candidate protein facilitating strand separation in *Agrobacterium* plasmids is virC1. Because the overdrive sequence has been shown to function as binding site for virC1 (Toro et al., 1989), it is

possible that virC1 also binds to the broader sequence context of the various AC-rich regions. The function of virC1 has not been elucidated yet, but homology with bacterial parA partitioning proteins suggests that virC1 may bind single-stranded DNA to extend unwound DNA regions (Easter and Gober, 2002).

Apart from the presence of potentially DNA unwinding sequences within extended right border regions, the AT-rich sequences upstream from left borders and left border alternatives represented low helical stability regions (C.M. Rommens, unpublished data). Thus, the facilitation of local DNA unwinding may not just be important for initiation but also for termination of DNA transfer in a similar way as described for conjugative DNA transfer (Ryan et al., 2004). In the case of pTiC58, this AT-rich DNA contains the sequence 5'-TTCAGTACATTA, which resembles the putative binding site for integration host factor (IHF) binding, 5'-[A/T]-TCAANNNTTA (Friedman, 1988; McGuire et al., 2000). Because IHF is involved in virtually all forms of nucleoid manipulation including DNA unwinding (Ryan et al., 2004), it is possible that this protein supports DNA cleavage at left borders in a similar way as reported previously for oriT (Karl et al., 2001).

*Agrobacterium* left border regions can also be characterized by the presence of at least four C-clusters at the junction between the borders and upstream DNA. The C-clusters may form into tertiary quadruplexes at slightly acid or neutral pH, in a similar manner as described for mammalian C-clusters (Neidle and Parkinson, 2003; Zarudnaya et al., 2003). It is possible that the complex folding associated with C-clusters negatively affects processes that are associated with DNA transfer, possibly limiting the frequency of backbone DNA transfer. The organization of extended regions for the initiation and termination of DNA transfer is summarized in Figure 5, accounting for domains of local DNA context around both the primary and secondary cutting sites described in this work. These sequence requirements should be considered in efforts to create novel all-native P-DNA vectors for

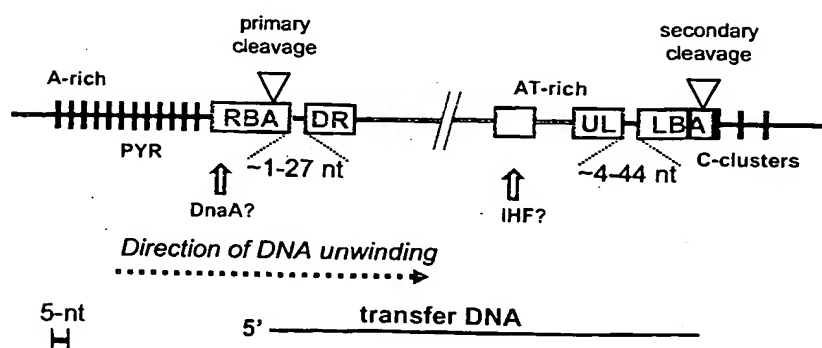


Figure 5. General organization of extended border regions. Putative sites for DnaA and IHF are indicated with open vertical arrows. Right and left borders are shown as yellow boxes. The positions of cleavage sites are depicted with triangles. The direction of DNA unwinding is indicated with a dashed horizontal arrow.

Agrobacterium- or Rhizobium-mediated plant transformation, en route to creating intragenically modified crops.

## MATERIALS AND METHODS

### PCR-Based Isolation of Right Border Alternatives

Plant DNAs (2 µg), partially digested with *SauIII*A, were ligated with 192-bp *Bam*HI-*Eco*RV fragments of pBR322. The resulting DNAs were used as templates for amplification with the primer pair 5'-YGR CAG GAT ATA TNN NNN KGT AAA C-3' (degenerate border) and 5'-GAC CAC ACC CGT CCT GTG-3' (anchor primer), with 49°C annealing temperature and 2.5-min extension time. Subsequent PCRs were performed with the amplified DNAs ligated with pGEM-T as templates using the border primer together with either SP6 or T7 primers at a slightly higher annealing temperature (52°C). The products of these reactions were inserted into pGEM-T and sequenced to design primers for conventional inverse PCRs to determine the actual sequence of the putative right border alternative.

### Database Searches and Sequence Analysis

Publicly available databases, including those maintained by the National Center For Biotechnology Information, were searched for T-DNA border sequences using the Motif Alignment and Search Tool (Bailey and Gribskov, 1998) and advanced BLASTN with penalty for nucleotide mismatch set at -1 and expect value at  $10^5$  (Altschul et al., 1997). Additional databases that were searched include those covering Solanaceae (<http://www.sgn.cornell.edu/>), Compositae (<http://compositdb.ucdavis.edu/>), and *Medicago truncatula* (<http://www.genome.ou.edu/medicago.html>). Folding of single-stranded DNA was studied using both the Mfold Web server (Zuker, 2003) and the Bielefeld Bioinformatics Server (Giegerich et al., 2004). Free energy profiles for unwinding 50-bp windows of double-stranded DNA across Ti plasmids were estimated using WEB-THERMODYN (Huang and Kowalski, 2003).

### Plasmid Construction

The plasmid that was used for construction of pSIM-T vectors comprised unique *Kpn*I and *Spe*I sites flanked by sequences upstream from the right border of pTi15955. Details are described elsewhere (Rommens et al., 2004). Both right borders and right border alternatives were inserted into this plasmid by ligating linearized plasmid DNA with annealed primers comprising the sequences of interest flanked by the sticky ends of *Kpn*I and *Spe*I, respectively.

### Isolation of Plant DNA Fragments

The following primers were used to amplify DNA fragments from plants: 5'-TTA TGC GGG CTA ACG GCG TG (forward) and 5'-GGG CCC GGT ACC CGT TAG GGC TAG (reverse) for the alfalfa (*Medicago sativa*) DNA fragment downstream from Ms01 of pSIM843; 5'-GGG CCC GGT ACC CGG AGG AGA CTC (forward) and 5'-GAA GAA GCA CAC CGG CAC TGG AAT T (reverse) for the potato (*Solanum tuberosum*) DNA fragment downstream from St02 of pSIM781; 5'-GAG GTA TAG AGG CAT GTC TGG CGT GAT C (forward) and 5'-GTG AAG TTT ATA ACA TGT TGA AGG AGC TCC A (reverse) for the potato DNA fragment following St02 of pSIM793; 5'-GGT ACC CTC TGT TGA CCA GGA TAT G (forward) and 5'-ACT AGT CAT TAC CAA CAA ATA TAT CCT GG (reverse) for the tomato (*Lycopersicon esculentum*) DNA fragment downstream from Le01 of pSIM582; 5'-CTT ACG TAG AAT TCT GTG CCA TG (forward) and 5'-GTT GGA TCT AGT TCT TGA AGT T (reverse) for the potato DNA fragment upstream from the second St01 copy of pSIM781B; and 5'-AAT TTT ATA TTA CAT TTA TAT TTG TG (forward) and 5'-CAT AAC AAA AAA AAA TTC TAT AAA TTA T (reverse) for the alfalfa DNA fragment preceding the second copy of the Ms01 element of pSIM843B.

### Plant Transformation

Binary vectors were introduced into *Agrobacterium tumefaciens* LBA4404 cells as follows. Competent LBA4404 cells (50 µL) were incubated for 5 min on ice in the presence of 1 µg of vector DNA, frozen for about 15 s in liquid nitrogen, and incubated at 37°C for 5 min. After adding 1 mL of liquid broth,

the treated cells were grown for 3 h at 28°C and plated on liquid broth/agar containing streptomycin (100 mg/L) and kanamycin (100 mg/L). The vector DNAs were then isolated from overnight cultures of individual LBA4404 colonies and examined by restriction analysis to confirm the presence of intact plasmid DNA.

A 10-fold dilution of an overnight-grown *Agrobacterium* culture was grown for 4 to 5 h, precipitated for 15 min at 3,800 rpm, washed with Murashige and Skoog liquid medium (PhytoTechnology, Shawnee Mission, KS) supplemented with Suc (3%, pH 5.7), and resuspended in the same medium to an optical density at 600 nm of 0.2 (for evaluation of new borders using pSIM-T vectors) or 0.04 (to assess the efficacy of new border-flanking DNA sequences). The suspension was then used to infect leaf explants of 3-week-old in vitro grown tobacco (*Nicotiana tabacum*) plants. Infected tobacco explants were incubated for 2 d on coculture medium (one-tenth Murashige and Skoog salts, 3% Suc, pH 5.7) containing 6 g/L agar at 25°C in a Percival growth chamber (16-h-light photoperiod) and subsequently transferred to M401/agar (PhytoTechnology) medium containing timentin (150 mg/L) and kanamycin (100 mg/L). Two weeks after inoculation, the number of calli per leaf explants was scored, and averages of all the explants for each of the treatments were calculated. The tobacco calli assay is highly quantitative because individual calli can be counted, and the large sample size (at least 50 different explants/infection) permits statistical analysis.

### Plant Genotyping

Plant DNA isolation and subsequent PCR were carried out as described previously (Rommens et al., 2004).

## ACKNOWLEDGMENTS

This article is dedicated to Dr. Kenneth Barton. We thank Rachel Perry, Lynda Zhang, Kristine Barney, and Artesia Stivison for excellent research assistance. We are grateful to Scott Simplot and Bill Whitacre for fruitful discussion and support.

Received July 22, 2005; revised July 22, 2005; accepted August 24, 2005; published October 21, 2005.

## LITERATURE CITED

- Altschul SE, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402
- Bailey TL, Gribskov M (1998) Methods and statistics for combining motif match scores. *J Comput Biol* 5: 211-221
- Betteridge T, Yang J, Pittard AJ, Praszkiar J (2004) Role of RepA and DnaA proteins in the opening of the origin of DNA replication of an IncB plasmid. *J Bacteriol* 186: 3785-3793
- Casas-Flores S, Rosales-Saavedra T, Herrera-Estrella A (2004) Three decades of fungal transformation: novel technologies. *Methods Mol Biol* 267: 315-325
- Easter J, Gober JW (2002) ParB-stimulated nucleotide exchange regulates a switch in functionally distinct ParA activities. *Mol Cell* 10: 427-434
- Friedman DI (1988) Integration host factor: a protein for all reasons. *Cell* 55: 545-554
- Garbarino JE, Belknap WR (1994) Isolation of a ubiquitin-ribosomal protein gene (*ubi3*) from potato and expression of its promoter in transgenic plants. *Plant Mol Biol* 24: 119-127
- Giegerich R, Voss B, Rehmsmeier M (2004) Abstract shapes of RNA. *Nucleic Acids Res* 32: 4843-4851
- Hansen G, Tempe J, Brevet J (1992) A T-DNA transfer stimulator sequence in the vicinity of the right border of pRi8196. *Plant Mol Biol* 20: 113-122
- Huang S, Gilbertson LA, Adams TH, Malloy KP, Reisenbigler EK, Birr DH, Snyder MW, Zhang Q, Luethy MH (2004) Generation of marker-free transgenic maize by regular two-border *Agrobacterium* transformation vectors. *Transgenic Res* 13: 451-461
- Huang Y, Kowalski D (2003) WEB-THERMODYN: sequence analysis software for profiling DNA helical stability. *Nucleic Acids Res* 31: 3819-3821

- Jakimowicz D, Chater K, Zakrzewska-Czerwinska J (2002) The ParB protein of *Streptomyces coelicolor* A3(2) recognizes a cluster of parS sequences within the origin-proximal region of the linear chromosome. *Mol Microbiol* 45: 1365–1377
- Jouanin L, Tourneur J, Tourneur C, Casse-Delbart F (1986) Restriction maps and homologies of the three plasmids of *Agrobacterium rhizogenes* strain A4. *Plasmid* 16: 124–134
- Karl W, Bamberger M, Zechner EL (2001) Transfer protein TraY of plasmid R1 stimulates Tral-catalyzed oriT cleavage in vivo. *J Bacteriol* 183: 909–914
- Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J* 11: 945–957
- Kunik T, Tzfira T, Kapulnik Y, Gafni Y, Dingwall C, Citovsky V (2001) Genetic transformation of HeLa cells by *Agrobacterium*. *Proc Natl Acad Sci USA* 98: 1871–1876
- Lessl M, Lanka E (1994) Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. *Cell* 77: 321–324
- Liu HX, Zhang M, Krainer AR (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev* 12: 1998–2012
- Long M, Rosenberg C (2000) Testing the "proto-splice sites" model of intron origin: evidence from analysis of intron phase correlations. *Mol Biol Evol* 17: 1789–1796
- McGuire AM, Hughes JD, Church GM (2000) Conservation of DNA regulatory motifs and discovery of new motifs in microbial genomes. *Genome Res* 10: 744–757
- Natale DA, Umek RM, Kowalski D (1993) Ease of DNA unwinding is a conserved property of yeast replication origins. *Nucleic Acids Res* 21: 555–560
- Neidle S, Parkinson GN (2003) The structure of telomeric DNA. *Curr Opin Struct Biol* 13: 275–283
- Nielsen KM (2003) Transgenic organisms—time for conceptual diversification? *Nat Biotechnol* 21: 227–228
- Patterson DJ, Yasuhara K, Ruzzo WL (2002) Pre-mRNA secondary structure prediction aids splice site prediction. In AK Dunker, RB Altman, L Hunter, TE Klein, eds, *Pacific Symposium on Biocomputing*. World Scientific Press, Singapore, pp 223–234
- Reddy AS (2004) Plant serine/arginine-rich proteins and their role in pre-mRNA splicing. *Trends Plant Sci* 9: 541–547
- Rommens CM (2004) All-native DNA transformation: a new approach to plant genetic engineering. *Trends Plant Sci* 9: 457–464
- Rommens CM, Humara JM, Ye J, Yan H, Richael C, Zhang L, Perry R, Swords K (2004) Crop improvement through modification of the plant's own genome. *Plant Physiol* 135: 421–431
- Roth A, Messer W (1998) High-affinity binding sites for the initiator protein DnaA on the chromosome of *Escherichia coli*. *Mol Microbiol* 28: 395–401
- Ryan VT, Grimwade JE, Camara JE, Crooke E, Leonard AC (2004) *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Mol Microbiol* 51: 1347–1359
- Scheiffele P, Pansegrau W, Lanka E (1995) Initiation of *Agrobacterium tumefaciens* T-DNA processing. Purified proteins VirD1 and VirD2 catalyze site- and strand-specific cleavage of superhelical T-border DNA in vitro. *J Biol Chem* 270: 1269–1276
- Shurvinton CE, Hodges L, Ream W (1992) A nuclear localization signal and the C-terminal  $\Omega$  sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. *Proc Natl Acad Sci USA* 89: 11837–11841
- Shurvinton CE, Ream W (1991) Stimulation of *Agrobacterium tumefaciens* T-DNA transfer by overdrive depends on a flanking sequence but not on helical position with respect to the border repeat. *J Bacteriol* 173: 5558–5563
- Speck C, Messer W (2001) Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *EMBO J* 20: 1469–1476
- Szybalski W, Kubinski H, Sheldrick P (1966) Pyrimidine clusters on the transcribing strand of DNA and their possible role in the initiation of RNA synthesis. *Cold Spring Harb Symp Quant Biol* 31: 123–127
- Toro N, Datta A, Carmi OA, Young C, Prusti RK, Nester EW (1989) The *Agrobacterium tumefaciens* virC1 gene product binds to overdrive, a T-DNA transfer enhancer. *J Bacteriol* 171: 6845–6849
- van Haaren MJ, Sedee NJ, Schilperoort RA, Hooykaas PJ (1987) Overdrive is a T-region transfer enhancer which stimulates T-strand production in *Agrobacterium tumefaciens*. *Nucleic Acids Res* 15: 8983–8997
- Wang K, Herrera-Estrella L, Van Montagu M, Zambryski P (1984) Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell* 38: 455–462
- Zarudnaya MI, Kolomiets IM, Potyahaylo AL, Hovorun DM (2003) Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. *Nucleic Acids Res* 31: 1375–1386
- Ziemienowicz A, Merkle T, Schoumacher F, Hohn B, Rossi L (2001) Import of *Agrobacterium* T-DNA into plant nuclei: two distinct functions of VirD2 and VirE2 proteins. *Plant Cell* 13: 369–383
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31: 3406–3415
- Zupan JR, Zambryski P (1995) Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiol* 107: 1041–1047

### Exhibit 3

Example of BLAST results for the query sequence CGACAGGATATATTTGTTGGTAATC that represents the border motif YGRYAGGATATATWSNVBKGTAAWY. The border-like sequences are underlined.

Database: Database of GenBank+EMBL+DDBJ sequences from EST

>gi|25018487|gb|CA517930.1| KS10002H08 KS10 Capsicum annuum cDNA, mRNA sequence.

Query 4 CAGGATATATTTGTTGGTAAT 24  
|||||  
Sbjct 298 CAGGATATATTTGTTGGTAAT 318

CGAGATTATCTTCCAATTTTCTGCGATCAGAGTGGGAACTAAATCCTATGGTGATCAAAATGCTGTGTCTGAGAACGGAAATACTAATGGAG  
GCTTGCATTTTGACGACCTTGTGGATGATGATTTGGTCGAATCAGTGGAGCTCAGAACAATGAAATTTCCAGAAATCCATCTGCCTCTTAA  
GAGGTCAATATTAAGGGCTTTAGATCAATGGATAAGGAACCTGAAGTTACACCCCTCAATTGACTGTTTCTGCAGCGGGTCAACTGCTGTACT  
TTGGTAATGCAGGGCCAGGATATATTTGTTGGTAATGTTGGCGACTCGAGAGCTGTTTTGGCAACAAGAGATAAAGACGACAACCTAATGGCCG  
TGCAGTTGACAGTAGATTTGAAGCCTAATCTTCCAAGAGAAGCTGCTAGAATCCATAAGTGTAAGGAAGGGTTTTTGCCTACAAGACGAAC  
AGAGGTTGCACGTGTGTGGTTGCCAAAC

>gi|16248687|gb|BI934215.1| EST554104 tomato flower, anthesis Lycopersicon esculentum cDNA  
clone cTOD19K21 5' end, mRNA sequence.

Query 4 CAGGATATATTTGTTGGTAAT 24  
|||||  
Sbjct 413 CAGGATATATTTGTTGGTAAT 433

AACCTATGTGGAGTATTTGATGGGCATGGTCTCATGGTCATATGGTTGCCAGTAAAGTTCGAGATTATCTTCCAATCTTGCTGCGATCAGAA  
TGGGAACTAAATCTGTAGTGATCTAAATGATGTATCCAAGATGAAATACTAATGGAGGCTTACATTTTGACGACATTGTGGATGATGATT  
TGTTGAATCAGTGGAGATGAGAACAATGTACAATCCCAGAAATCCATCTACATATTAAGAGGTCAATATTAAGGGCTTTAGATCAATGGA  
TAAGGAACCTGAAGTTACACCCCTCAATTGACTGTTTCTGCAGCGGATCAACTGCTGTTACTTTGGTAAAGCAGGATATGATGTTGAAACATTAT  
ACTTTCAGAAAATAATATCAATTGAGTTGGAAGGGCCAGGATATATTTGTTGGTAATGTTGGTGACTCGAGAGCAGTTTGGCAACAAGAGATA  
AAGACAACCTCTTAATGGCTGTACAGTTGACAGTAGATTTGAAGCCTAATCTTCCAAGTCTGTTTGTCTCTTCTCTTACCCGCCCCACCCAAC  
ACAGTGCACATTAGCTGAAGCTGCACTTATATTTATTTCAATTGTTGAACTCAGTAGTTTCTGTCTTCTGCTTGTAGGAGAAGCTGCTAGAAT  
CCATAAATGTAAAGGAAGGGGTTTTGCACTGCAAGATGA

>gi|78085038|gb|DV513431.1| ZM\_BFb0194J20.f ZM\_BFb Zea mays cDNA 3', mRNA sequence.

Query 4 CAGGATATATTTGTTGGTA 22  
|||||  
Sbjct 178 CAGGATATATTTGTTAGTA 196

AAACGGAACACGACGTGTGTACAATTTTGAATCCTTCATACACCAACACACCTGGCAAAACCTCATGGAAGGAAGGAAATCATCAACCCAA  
ATGACTGAAAAATTGGGAGGCACTGGTCAAGCCTTCTACTCTGCTCTACTAGTCTAAGCTGCTGCTGCTGGCACCTTCAGCAGGATATAT  
TGTTAGTACTTAGTACGGTACGGTGGTGTCTCTTCTAGTCACTTGTCTCTGTTGGTGGTGTGGGCTGTGGCGGTGGCGGCGGCTGCTCTGCTCG  
TTTGGTCTCTGCTCTTGAGGGCAAGCAGCATGCTGCGGAGGTGTTATTGGTCATCATGCCAGGCGCAATTGACACTGACAAGGTTAGTACGA  
GCTCCGTCATCTCTCACACATCCCTACTGTGCGGATCGGACGGCTCAGACGGTCCGCGCCGAAGACCACGGGCGCCCGTACTGCGGGTGGA  
GGAGCATCACCACTGCGGCGGTGGAGGTAGCTGGGCCACACCCGGAAGTCCAACCGCTGGAGGACGATGGCCAGCGTGAGCTTACACTCCAG  
GAGCGCCAGGTTCTGGCCGATGCACATCCCCCGCCGAGCC

>gi|73933229|gb|DT516539.1| WS02431.B21\_L08 PTxD-ICC-N-A-14 Populus trichocarpa x Populus  
deltoides cDNA clone WS02431\_L08 3', mRNA sequence.

Query 4 CAGGATATATTTGTTGGTAAT 24  
|||||  
Sbjct 404 CAGGATATATGTGTCGGTAAT 384

TTTTTTTTTTTTTTTTTATCGCAATTGTGATTAAATTAATAATCATTGTGATAGAATCAAAATCCTTCCATGGATGCATATTAATCAAATGAA  
TTAATAAAGACTTGGAGCAAGCATGCAAGATACACAATGCATGAATTAAGCGAGATCCCTCCACATATTCGGTTGGAGGCAGGTGTCTGGCCC  
CACCGGAAGCGACAGCGCGGCAACAACGACCTCTAAATCTCTTGGTAGGTCCGCGTAACATAAGTCTGTCCCTCTCAGTCTCCTCCACGC  
CTCCCTCTCTTTACCTCTATATAACAAGCCCTTGAGGATCTGATGTACCAATCAAAACAAGATCGCACGGCACTTCATCATTCTCACGAAGCC  
ACACTAAATTACCGACACATATATCTGCGCTTGAATATGTTTCTTGATCCCTTCTTAACAATCCAACTTGCTTCTCATTGCGCTTCTTATC  
AGACAAATATCTATTATAATCATCCCAAGCCTCTTTCGTCGCAGACACAGCAAATATAAAATCAACGGCCCCCAAGTGCTCGCGGATTACCC  
GGGGTAATAAGCGACCATTAATTGAAGACACGAATCAACAAGAAATACTGGTTCATAAACCAGTGAAGTCTGCTCCATGAGATTTTAGGGAGGA  
AATTGAGGAGAGTGATTTCCGATTGGAAATGCGGTTGTCGCAGTAAAGATCATGTGTAGGCGATGACTCATGCTCGTTGATGTAGACGAACCG  
CTTCATTTCTTTGATTGACTTTGATTGATTGATTACATCATCACTCCAGACGACTACGCATTGCTTACCGT

**Exhibit 4**

Example of the isolation of a border-like sequence from potato DNA using the two-step PCR method described in Example 1.

| Step | Reference in<br>USN<br>10/607,538 | Sequence                              |  |
|------|-----------------------------------|---------------------------------------|--|
| 1    | Paragraph 333                     | 1 <sup>st</sup> border primer         | YGRCAGGATATATNNNNKGTAAAC   |
|      | Paragraph 333                     | Anchor primer                         | GACCACACCCGTCCTGTG   |
|      | Paragraph 335                     | 1 <sup>st</sup> amplified product     | CAGGATATATGTGAGGGTAAACAAATAACAACATAACATGGA<br>AATAGATATACTGAAACATAACTAAAAATTGAACTAGGT<br>TCAATAACCAAGTACTAATCATGAAGATAACATGCAGAA<br>CTGAATACTGAAATAAATGTTGAATACATGGTCAATGCA<br>TCATGAATCTGACTGTGGGAGCTACTATTAACATGACATA<br>AACCATGTGAGATAAACACGGAATCTGATGTATATTCCC<br>TATCGAGAGGACCTAATGTACCTTGTCGGGGTATAAAAT<br>CTATGGCGTGATCCACAGGACGGGTGTGGTC |
|      | Paragraph 336                     | Expected border-like sequence         | CAGGATATATGTGAGGGTAAAC   |
|      | Paragraph 337                     | 2 reverse primers                     | CAGTATTCAGTTCTGCATGTTATC (1 <sup>st</sup> ) and<br>GTACTTGGTTATTGAACCTAGTTC (2 <sup>nd</sup> )   |
| 2    | Paragraph 337                     | 2 direct primers                      | CTGACTGTGGGAGCTACTATTAAC (1 <sup>st</sup> ) and<br>GATGTATATTCCCTATCGAGAGGAC (2 <sup>nd</sup> )  |
|      |                                   | Restriction site                      | ATGCAT (NsiI)  |
|      |                                   | iPCR product (up to blunt-ended site) | CTTGTTATTGAACCTAGTTCAATTTTTAGTTATGTTTCAGT<br>ATATCTATTTCCATGTTTAGTTGTTATTTGTTTAGCTTA<br>GCATATATCCTGCATGC  |
|      | Paragraph 337                     | Actual border-like sequence           | GTTTAGCTTAGCATATATCCTGCAT  |
|      | Paragraph 339-346                 | Activity as border                    | Yes, with 89.7% of the activity of a conventional T-DNA border (reported 4/24/04)  |

Exhibit 5

| Plant | Sequence                   | Vector<br>pSIM | Core* | Efficiency<br>% T-DNA | Method**    |
|-------|----------------------------|----------------|-------|-----------------------|-------------|
|       | ANGATNTATNNNNNNGT          |                |       |                       |             |
| At    | TGACAGGATATATCGTGATGTCAAC  | 564            | I     | 64                    | AL138652    |
| St    | ATGCAGGATATATGCTAAGCTAAAC  | 605            | H     | 60                    | PCR         |
| St    | AGCCAGGATATATTTGCAGCTAGAA  | 736            | H     | 0.4                   | SGN-U180201 |
| St    | AGGCAGGATATATTTTGGGGTAGGT  | 737            | I     | 43                    | SGN-U179068 |
| St    | GGCCAGGATATATTTGTTGTAATG   | 551            | I     | 100                   | PCR         |
| St    | AGGCAGGATATATTTTGGGGTAAAC  | 565            | I     | 86                    | PCR         |
| St    | GGACAGGATATATAAAGTGTAAC    | 566            | H     | 6.6                   | PCR         |
| Le    | GGGCAGGATATATACAAGTCTAAAC  | 603            | H     | 17                    | BI421743    |
| Le    | GGGCAGGATATATTTTGGGGTAGGT  | 732            | I     | 25                    | SGN U152635 |
| Le    | AGCCAGGATATATTTGCAGCTAGAA  | 719            | H     | .17                   | SGN U158307 |
| Le    | CCGCAGGATATATTCAGAGGTAGAG  | 560            | I     | 102                   | SGN U158307 |
| Le    | CCGCAGGATATATTCAGAGGTATGC  | 634            | I     | 98.5                  | SGN U160333 |
| Le    | TTACAGGATATATACAGTGTAAACG  |                | H     | 8                     | OB683       |
| Le    | TGGCAGGATATATGGCCAGGAACAA  | 720            | H     | 2.4                   | SGN-U155745 |
| Le    | ACGCAGGATATATCATCGGTACTT-  |                | H     | 5.5                   | AY013255    |
| Le    | GGGCAGGATATATTTTGGGGTAGGT  | 721            | I     | 30                    | SGN-U153074 |
| Le    | TTACAGGATATATGCAAAATAATTC  | 633            | H     | 6.5                   | SGN-U160800 |
| Le    | TGGCAGGATATATTTGATGCAATG-  | 724            | H     | .9                    | SGN-U162952 |
| Le    | TCACAGGATATATTAGACAGTTTCC  | 725            | I     | 35                    | SGN-U149980 |
| Le    | AAGCAGGATATATCTACATGGTACC  | 741            | H     | 1.1                   | AY013255    |
| Le    | TGGCAGGATATATTT-GATGCAATC  |                | H     | .19                   | AW031162    |
| Le    | GTACAGGATATATTGAGTTCTAAGA  | 550            | H     | 1.7                   | AW618843    |
| Le    | TGTCAGGATATATGTCAACCGGAAC  | 551            | H     | 0.1                   | AW398191    |
| Le    | GGCCAGGATATATTTGTTGTAATG   | 732            | I     | 100                   | SGN U152635 |
| Ms    | CGGCAGGAGGTATACAGAGGTATAC  | 844            | H     | 168                   | PCR         |
| Mt    | CGGCAGGAGGTATACAGAGGTATAC  | 568            | H     | 32                    | CA919279    |
| Ms    | CGGCAGGATGTATACAGAGGTATAC  |                | I     | 56                    | Modified    |
| Hv    | GGGCAGGATGTATCATTTGGTATAT  | 570            | I     | 90                    | TC137112    |
| Hv    | TGGCAGGAACATATGAGGCGGTAGTG | 571            | H     | 0.3                   |             |
| Hv    | GGACAGGATATATAAAGTGTAAC    | 566            | H     | 6.5                   |             |
| Bn    | CTACAGGATTTATTATACAGTAGGG  | 1064           | I     | 2.6                   | BnL58       |
| Bn    | TGACAGGACATATAGAACAGTTCCT  | 1065           | H     | 4.2                   | BnL59       |
| Zm    | CCACAGGATATATAAATGCGTAAGC  | 1063           | I     | 9                     | CF052264    |
| Ta    | GGACAGGATATATTGCTCCGTATAT  | 928            | I     | 78                    | Ca500064    |
| Os    | AGCCAGGATATATCCCTTGAGTAAGT | 923            | I     | 61                    | AL772425    |
| Os    | ATGCAGGATATATTTTTTTGTAGTG  | 924            | I     | 51                    | AP004161    |
| Os    | ACACAGGATATATATATACGTACAT  | 925            | I     | 44                    | AP005498    |
|       |                            |                |       |                       |             |
|       |                            |                |       |                       |             |
|       |                            |                |       |                       |             |
|       |                            |                |       |                       |             |

\*: core sequence = ANGATNTATNNNNNNGT (I = identical; H = homologous)

\*\*:: method: Sequences isolated by using PCR-based methods are indicated with "PCR"; sequences identified in databases are shown with their accession number.